

PCT WELTORGANISATION FÜR GEISTIGES EIGENTUM
Internationales Büro
INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE

| 1 1 | (11) Internationale Veröffentlichungsnummer: WO 00/1723 |
|--------------------------------------|--|
| A2 | (43) Internationales Veröffentlichungsdatum: 30. März 2000 (30.03.00 |
| nber 199 (22.09.99 .98) D D | BG, BR, BY, CA, CH, CN, CR, CU, CZ, DK, DM, E ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, J KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MI MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SI SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, U UZ, VN, YU, ZA, ZW, ARIPO Patent (GH, GM, KE, L MW, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AK AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Pate (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, I LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CC, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). |
| | YNTHETIC PATHWAY OSYNTHESEWEGS |
| | A2 299/0705: nber 1999 (22.09.99 DI (DE/DE) thel, Park |

5-phosphat- Synthase- Gen und das gcpE-Gen des 1-Desoxy- D-xylulose- Biosynthesewegs und ihre Verwendung zur Transformation von Vektoren, Wirtsorganismen und Pflanzen und zur Bestimmung von Stoffen, die diesen Biosyntheseweg inhibieren.

LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

| AL | Albanien | ES | Spanien | LS | Lesotho | SI | Slowenien |
|----|------------------------------|----------|-----------------------------|----|-----------------------------|----|------------------------|
| AM | Armenien | E3 Fi | Finnland | LT | Litauen | SK | Slowakei |
| AT | Österreich | FR | | LU | | SN | |
| | | | Frankreich | | Luxemburg | - | Senegal |
| AU | Australien | GA | Gabun | LV | Lettland | SZ | Swasiland |
| ΑZ | Aserbaidschan | GB | Vereinigtes Königreich | MC | Monaco | TD | Tschad |
| BA | Bosnien-Herzegowina | GE | Georgien | MD | Republik Moldau | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagaskar | TJ | Tadschikistan |
| BE | Belgien | GN | Guinea | MK | Die ehemalige jugoslawische | TM | Turkmenistan |
| BF | Burkina Faso | GR | Griechenland | | Republik Mazedonien | TR | Türkei |
| BG | Bulgarien | HU | Ungarn | ML | Mali | TT | Trinidad und Tobago |
| BJ | Benin | IE | Irland | MN | Mongolei | UA | Ukraine |
| BR | Brasilien | IL | Israel | MR | Mauretanien | UG | Uganda |
| BY | Belarus | IS | Island | MW | Malawi | US | Vereinigte Staaten von |
| CA | Kanada | IT | Italien | MX | Mexiko | | Amerika |
| CF | Zentralafrikanische Republik | JP | Japan | NE | Niger | UZ | Usbekistan |
| CG | Kongo | KE | Kenia | NL | Niederlande | VN | Vietnam |
| CH | Schweiz | KG | Kirgisistan | NO | Norwegen | YU | Jugoslawien |
| CI | Côte d'Ivoire | KP | Demokratische Volksrepublik | NZ | Neuseeland | zw | Zimbabwe |
| CM | Kamerun | | Korea | PL | Polen | | |
| CN | China | KR | Republik Korea | PT | Portugal | | |
| CU | Kuba | KZ | Kasachstan | RO | Rumanien | | |
| CZ | Tschechische Republik | LC | St. Lucia | RU | Russische Föderation | | |
| DE | Deutschland | LI | Liechtenstein | SD | Sudan | | |
| DK | Dänemark | LK | Sri Lanka | SE | Schweden | | |
| EE | Estland | LR | Liberia | SG | Singapur | | |
| | | | | | - | | |

Gene des 1-Desoxy-D-xylulose-Biosynthesewegs

Die vorliegende Erfindung betrifft DNA-Sequenzen, die bei Integration in das Genom von Viren, Eukaryonten und Prokaryonten die Isoprenoid-Biosynthese verändern sowie gentechnologische Verfahren zur Herstellung dieser transgenen Viren, Eukaryonten und Prokaryonten. Außerdem betrifft sie Verfahren zur Identifiziereung von Stoffen mit herbizider, antimikrobieller, antiparasitärer, antiviraler, fungizider, bakterizider Wirkung bei Pflanzen und antimikrobieller, antiparasitärer, antimykotischer, antibakterieller und antiviraler Wirkung bei Mensch und Tier.

Der Biosyntheseweg zur Bildung von Isoprenoiden über den klassischen Acetat/ Mevalonat-Weg und einen alternativen, Mevalonat-unabhängigen Biosyntheseweg, den Desoxy-D-xylulose-Phosphat-Weg, ist bereits bekannt (Rohmer, M., Knani, M., Simonin, P., Sutter, B., and Sahm, H. (1993): Biochem. J. 295: 517-524).

Es ist aber nicht bekannt, wie und über welche Wege in Viren, Eukaryonten und Prokaryonten eine Änderung der Isoprenoidkonzentration über den Desoxy-D-xylulose-Phoshat-Weg erreicht werden kann. In Fig. 1 ist dieser Biosyntheseweg dargestellt.

Es werden daher DNA-Sequenzen zur Verfügung gestellt, die für die 1-Desoxy-D-xylulose-5-phosphat-Synthase (DOXP-Synthase),: 1-Desoxy-D-xylulose-5-phosphatreduktoisomerase(DOXP-Reduktoisomerase) oder das gcpE-Protein kodieren. Alle drei Gene und Enzyme sind an der Isoprenoid-Biosynthese beteiligt.

Das gcpE-Protein hat eine Kinasefunktion und katalysiert die Phosphorylierung eines Zuckers oder eines Phosphorzuckers oder einer Vorstufe der Isoprenoidbiosynthese, insbesondere die Phosphorylierung von 2-C-Methyl-D-erythritol, 2-C-Methyl-D-erythritol-phosphat, insbesondere 2-C-Methyl-D-erythritol-4-phosphat, 2-C-Methyl-D-erythrose, 2-C-Methyl-D-erythrose-

phosphat, insbesondere 2-C-Methyl-D-erythrose-4-phosphat. In der Vorstufe der Isoprenoidsynthese katalysiert das gcpE-Protein insbesondere die Phosphorylierung der folgenden Substanzen:

 $\begin{array}{l} \text{CH}_2\left(\text{OH}\right) - \text{C}\left(\text{CH}_3\right) = \text{C}\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_2\left(\text{OH}\right) - \text{C}\left(\text{CH}_3\right) = \text{C}\left(\text{OH}\right) - \text{CH}_2 - \text{OH}, \\ \text{CH}_2\left(\text{OH}\right) - \text{CH}\left(\text{CH}_3\right) - \text{CO} - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_2 = \text{C}\left(\text{CH}_3\right) - \text{CO} - \text{CH}_2 - \text{OH} \\ \text{CH}_2 = \text{C}\left(\text{CH}_3\right) - \text{CO} - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_2 = \text{C}\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{OH}, \\ \text{CH}_2 = \text{C}\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_2 = \text{C}\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{OH}, \\ \text{CH}_2\left(\text{OH}\right) - \text{C}\left(\text{CH}_2\right) - \text{C}\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_2\left(\text{OH}\right) - \text{C}\left(\text{CH}_3\right) - \text{C}\left(\text{OH}\right) - \text{CH}_2 - \text{OH}, \\ \text{CH}_2\left(\text{OH}\right) - \text{C}\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_2\left(\text{OH}\right) - \text{C}\left(\text{OH}\right) - \text{CH}_2 - \text{OH}, \\ \text{CH}_3\left(\text{OH}\right) - \text{C}\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}\left(\text{OH}\right) = \text{C}\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{OH}, \\ \text{CH}_3\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_3\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{H}, \\ \text{CH}_3\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_3\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{H}. \\ \end{array}$

Die DOXP-Synthase katalysiert die Kondensation von Pyruvat und Glyceraldehyd-3-phosphat zu 1-Deoxy-D-xylulose-5-phosphat und die DOXP-Reduktoisomerase katalysiert die Umwandlung von 1-Deoxy-D-xylulose-5-phosphat zu 2-C-Methyl-D-erythritol-4-phosphat. (siehe Fig. 1).

Die Erfindung betrifft die folgenden DNA-Sequenzen:
DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO:2
dargestellten Aminosäuresequenz codieren oder für ein Analoges
oder Derivat des Polypeptids gemäß SEQ ID NO:2, worin eine oder
mehrere Aminosäuren deletiert, hinzugefügt oder durch andere
Aminosäuren substituiert worden sind,

DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO:4 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:4, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind,

sowie DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 6 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 6, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind.

Die Gene und ihre Genprodukte (Polypeptide) sind im Sequ nzprotokoll mit ihrer Primärstruktur aufgeführt und haben folgende Zuordnung:

SEQ ID NO:1: 1-Desoxy-D-xylulose-5-phosphatreduktoisomerase-Gen

SEQ ID NO:2: 1-Desoxy-D-xylulose-5-phosphatreduktoisomerase

SEQ ID NO:3: 1-Desoxy-D-xylulose-5-phosphat-Synthase-Gen

SEQ ID NO:4: 1-Desoxy-D-xylulose-5-phosphat-Synthase

SEQ ID NO:5: gcpE-Gen

SEQ ID NO:6 : gcpE-Proteine.

Die DNA-Sequenzen stammen alle aus Plasmodium falciparum.

Außer den im Sequenzprotokoll genannten DNA-Sequenzen sind auch solche geeignet, die infolge der Degeneration des genetischen Codes eine andere DNA-Sequenz besitzen, jedoch für das gleiche Polypeptid oder für ein Analoges oder Derivat des Polypeptids kodieren, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind.

Die erfindungsgemäßen Sequenzen eignen sich für die Expression von Genen in Viren, Eukaryonten und Prokaryonten, die für die Isoprenoid-Biosynthese des 1-Desoxy-D-xylulose-Wegs verantwortlich sind.

Erfindungsgemäß gehören zu den Eukaryonten oder eukaryontischen Zellen tierischen Zellen, Pflanzenzellen, Algen, Hefen, Pilze und zu den Prokaryonten oder prokaryontischen Bakterien Archaebakterien und Eubakterien.

Bei Integration einer DNA-Sequenz in ein Genom, auf der eine der oben angegebenen DNA-Sequenzen lokalisiert ist, wird die Expression der oben beschriebenen Gene in Viren, Eukaryonten und Prokaryonten ermöglicht. Die erfindungsgemäß transformierten Viren, Eukaryonten und Prokaryonten werden in an sich bekannter Weise gezüchtet und das währenddessen gebildete Isoprenoid isoliert und gegebenenfalls gereinigt. Nicht alle Isoprenoide müssen isoliert werden, da die Isoprenoide in einigen Fällen direkt in die Raumluft abgegeben werden.

Die Erfindung betrifft ferner ein Verfahren zur Herstellung von transgenen Viren, Eukaryonten und Prokaryonten zur Veränderung

des Isoprenoid-Gehaltes, das die folgenden Schritte enthält.

- a) Herstellung einer DNA-Sequenz mit folgenden Teilsequenzen
 - i) Promotor, der in Viren, Eukaryonten und Prokaryonten aktiv ist und die Bildung einer RNA im vorgesehenen Zielgewebe oder den Zielzellen sicherstellt,
 - ii) DNA-Sequenz, die für ein Polypeptid mit der in SEQ ID NO:2,4 oder 6 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:2,4 oder 6,
 - iii) 5'- und 3'-nichttranslatierte Sequenz, die in Viren, Eukaryonten und Prokaryonten die Expression der bezeichneten Gene ermöglichen oder verbessern,
- b) Transfer und Einbau der DNA-Sequenz in das Genom von Viren, prokaryontischen oder eukaryontischen Zellen mit oder ohne Verwendung eines Vektors (z.B. Plasmid, virale DNA).

Aus derart transformierten Pflanzenzellen können die intakten ganzen Pflanzen regeneriert werden.

Die für die Proteine kodierenden Sequenzen mit den Nukleotidabfolgen Seg ID NO:1, Seg ID NO:3 und Seg ID NO: 5 können mit einem die Transkription in bestimmten Organen oder Zellen sicherstellenden Promotor versehen werden, der in sense-Orientierung (3'-Ende des Promotors zum 5'-Ende der kodierenden Sequenz) an die Sequenz, die das zu bildende Protein kodiert, gekoppelt ist. An das 3'-Ende der kodierenden Segeunz wird ein die Termination der mRNA-Synthese bestimmendes Terminationssignal angehängt. Um das zu exprimierende Protein in bestimmte subzelluläre Kompartimente, wie Chloroplasten, Amyloplasten, Mitochondrien, Vakuole, Cytosol oder Interzellularräume zu dirigieren, kann zwischen den Promotor und die kodierende Sequenz noch eine für eine sogenannte Signalsequenz oder ein Transitpeptid kodierende Sequenz gesetzt werden. In einigen Fällen ist es erforderlich, Sequenzen einzufügen, die für eine Signalsequenz am COOH-Terminus des Proteins kodieren. Die Sequenz muß im glei-

chen Leserahmen wie die kodierende Sequenz des Proteins sein. Zur Vorbereitung der Einführung der erfindungsgemäßen DNA-Sequenzen in höhere Pflanzen sind eine große Anzahl von Klonierungsvektoren verfügbar, die ein Replikationssignal für E.coli und einen Marker beinhalten, der eine Selektion der transformierten Zellen erlaubt. Je nach Einführungsmethode gewünschter Gene in die Pflanze können weitere DNA-Sequenzen erforderlich sein. Werden zum Beispiel für die Transformation der Pflanzenzelle das Ti- oder Ri-Plasmid verwendet, so muß mindestens eine rechte Begrenzung, häufig jedoch die rechte und die linke Begrenzung der Ti- und Ri-Plasmid T-DNA als Flankenbereich den einzuführenden Genen eingefügt werden. Die Verwendung von T-DNA für die Transformation von Pflanzenzellen ist intensiv untersucht und ausreichend in EP 120516; Hoekama, in: The Binary Plant Vector System, Offset-drukkerij Kanters B.V. Alblasserdam (1985), Chapter V; Fraley et al., Crit.Rev.Plant Sci. 4,1-46 und An et al. (1985) EMBO J. 4, 277-287 beschrieben worden. Ist die eingeführte DNA einmal im Genom integriert, so ist sie in der Regel stabil und bleibt auch in den Nachkommen der ursprünglich transformierten Zellen erhalten. Sie erhält normalerweise einen Selektionsmarker, der den transformierten Pflanzenzellen Resistenz gegenüber einem Biozid oder einem Antibiotikum, wie Kanamycin, G 418, Bleomycin, Hygromycin oder Phosphinotricin u.a. vermittelt. Der individuell verwendete Marker sollte daher die Selektion transformierter Zellen gegenüber Zellen, denen die eingefügte DNA fehlt, gestatten.

Für die Einführung von DNA in eine Pflanze stehen viele Techniken zur Verfügung. Diese Techniken umfassen die Transformation mit Hilfe von Agrobakterien, z.B. Agrobacterium tumefaciens, die Fusion von Protoplasten, die Mikroinjektion von DNA, die Elekroporation, sowie ballistische Methoden und die Virusinfektion. Aus dem transformierten Pflanzenmaterial können dann im geeigneten Medium, welches Antibiotika oder Biozide zur Selektion enthalten kann, wieder ganze Pflanzen regeneriert werden. Bei der Injektion und Elektroporation sind an sich keine speziellen Anforderungen an die Plasmide gestellt. Sollen aber aus derartig transformierten Zellen ganze Pflanzen regeneriert werden, ist die Anwesenheit eines selektierbaren Markergens not-

wendig. Die transformierten Zellen wachsen innerhalb der Pflanzen in der üblichen Weise (McCormick et al. (1986), Plant Cell Reports 5, 81-84). Die Pflanzen können normal angezogen werden und mit Pflanzen, die die gleiche transformierte Erbanlage oder andere Erbanlagen haben, gekreuzt werden. Die daraus entstehenden Individuen haben die entsprechenden phänotypischen Eigenschaften.

Weiterhin sind Gegenstand der Erfindung Expressionsvektoren, die eine oder mehrere der erfindungsgemäßen DNA-Sequenzen enthalten. Solche Expressionsvektoren erhält man, indem man die erfindungsgemäßen DNA-Sequenzen mit geeigneten funktionellen Regulationssignalen versieht. Solche Regulationssignale sind DNA-Sequenzen, die für die Expression verantwortlich sind, beispielsweise Promotoren, Operatoren, Enhancer, ribosomale Bindungsstellen, und die vom Wirtsorganismus erkannt werden.

Gegebenenfalls können noch weitere Regulationssignale, die beispielsweise Replikation oder Rekombination der rekombinanten DNA im Wirtsorganismus steuern, Bestandteil des Expressionsvektors sein.

Ebenso gehören die mit den erfindungsgemäßen DNA-Sequenzen oder Expressionsvektoren transformierten Wirtsorganismen zum Gegenstand der Erfindung.

Für die Expression der erfindungsgemäßen Enzyme eignen sich besonders solche Wirtszellen und Organismen, die keine intrinsischen Enzyme mit der Funktion der DOXP-Synthase, der DOXP-Reduktoisomerase oder des gcpE-Proteins aufweisen. Dies trifft für Archaebacterien, Tiere, Pilze, Schleimpilze und einige Eubakterien zu. Durch das Fehlen dieser intrinsischen Enzymaktivitäten wird die Detektion und Aufreinigung der rekombinanten Enzyme wesentlich erleichtert. Auch wird es erst dadurch möglich, mit geringem Aufwand die Aktivität und insbesondere die Hemmung der Aktivität der erfindungsgemäßen rekombinanten Enzyme durch verschiedenen Chemikalien und Pharmaka in Rohextrakten aus den Wirtszellen zu messen.

Die Expression der erfindungsgemäßen Enzyme erfolgt vorteilhafterweise dann in eukaryontischen Zellen, wenn posttranslatorische Modifikationen und eine native Faltung der Polypeptidkette erreicht werden soll. Außerdem wird in Abhängigkeit vom Expressionssystem bei der Expression genomischer DNA-Sequenzen erreicht, daß Introns durch Spleißen der DNA beseitigt und die Enzyme in der für die Parasiten charakteristischen Polypeptidsequenz produziert werden. Für Introns codierende Sequenzen können auch durch rekombinante DNA-Technologie aus den zu exprimierenden DNA-Sequenzen beseitigt oder experimentell eingefügt werden.

Die Isolierung des Proteins kann aus der Wirtszelle oder dem Kulturüberstand der Wirtszelle nach dem Fachmann bekannten Verfahren erfolgen. Es kann auch eine in vitro Reaktivierung der Enzyme erforderlich sein.

Zur Erleichterung der Aufreinigung können die erfindungsgemäßen Enzyme oder Teilsequenzen der Enzyme als Fusionsprotein mit verschiedenen Peptidketten exprimiert werden. Dazu eigenen sich besonders Oligo-Histidin-Sequenzen und Sequenzen, die von der Glutathion-S-Transferase, Thioredoxin oder Calmodulin-bindenden Peptiden abgeleitet sind.

Weiterhin können die erfindungsgemäßen Enzyme oder Teilsequenzen der Enzyme als Fusionsprotein mit solchen, dem Fachmann bekannten, Peptidketten exprimiert werden, daß die rekombinanten Enzyme in das extrazelluläre Millieu oder in bestimmte Kompartimente der Wirtszellen transportiert werden. Dadurch kann sowohl die Aufreinigung, als auch die Untersuchung der biologischen Aktivität der Enzyme erleichtert werden.

Bei der Expression der erfindungsgemäßen Enzyme kann es sich als zweckmäßig erweisen, einzelne Codone zu verändern. Dabei

ist der gezielte Austausch von Basen in der kodierenden Region auch sinnvoll, wenn die genutzten Codone in den Parasiten abweichend sind von der Codonnutzung im heterologen Expressionssystem, um eine optimale Synthese des Proteins zu gewährleisten.

Weiterhin können die erfindungsgemäßen Enzyme unter standardisierten Bedingungen durch dem Fachmann bekannte Techniken durch in vitro-Translation gewonnen werden. Dafür geeignete Systeme sind Kaninchen-Reticulozyten- und Weizenkeimextrakte und Bakterienlysate. Auch kann in vitro transskribierte mRNA in Xenopus-Oocyten translatiert werden.

Durch chemische Synthese können Oligo- und Polypeptide hergestellt werden, deren Sequenzen aus der Peptidsequenz der erfindungsgemäßen Enzyme abgeleitet sind. Bei geeigneter Wahl der Sequenzen besitzen derartige Peptide Eigenschaften, die für die erfindungsgemäßen Enzyme charakteristisch sind. Derartige Peptide können in großen Mengen hergestellt werden und eignen sich besonders für Studien über die Kinetik der Enzymaktivität, die Regulation der Enzymaktivität, die dreidimensionale Struktur der Enzyme, die Hemmung der Enzymaktivität durch verschiedenen Chemikalien und Pharmaka und die Bindungsgeometrie und Bindugnsaffinität verschiedener Liganden.

Vorzugsweise wird zur rekombinanten Herstellung der erfindungsgemäßen Enzyme eine DNA mit den Nukleotiden aus den Sequenzen SEQ ID NO: 1, 3 und 5 verwendet.

Die Erfindung umfaßt daher außerdem ein Verfahren zum Screening nach Verbindungen, die desDesoxy-D-xylulose-Phosphat-Stoffwechselweg inhibieren. Gemäß diesem Verfahren wird ein Wirtsorganismus, der einen rekombinanten Expressionsvektor enthält, wobei der Vektor zumindest einen Teil der Olignukleotidsequenz gemäß SEQ ID NO:1, SEQ ID NO: 3 oder SEQ ID NO: 5 oder Varianten oder Homologe dieser aufweist, und außerdem eine

Verbindung, von der vermutet wird, daß sie eine antimikrobielle, antiparasitäre, antibakterielle, antivirale und antimykotische Wirkung bei Mensch und Tier oder eine antimikrobielle, antivirale, bakterizide, herbizide oder fungizide Wirkung bei Pflanzen hat, bereitgestellt. Anschließend wird der Wirtsorganismus mit der Verbindung in Kontakt gebracht und die Wirksamkeit der Verbindung bestimmt.

Ein weiterer Gegenstand dieser Erfindung sind Methoden zur Bestimmung der enzymatische Aktivität des gcpE-Proteins. Diese kann nach bekannten Verfahren bestimmt werden. Hierbei wird die Phosphorylierung eines Zuckers oder eines Phosphorzuckers oder einer Vorstufe der Isoprenoidbiosynthese, insbesondere die Phosphorylierung von 2-C-Methyl-D-erythritol, 2-C-Methyl-D-erythritol-phosphat, insbesondere 2-C-Methyl-D-erythritol-4-phosphat, 2-C-Methyl-D-erythrose, 2-C-Methyl-D-erythrose-phosphat, insbesondere 2-C-Methyl-D-erythrose-4-phosphat, detektiert. Ein weiterer Gegenstand dieser Erfindung ist die Verwendung dieser Meßverfahren zur Ermittlung von Stoffen, die die Aktivität der jeweiligen Enzyme inhibieren.

Die enzymatische Aktivität von DOXP-Synthase und DOXP-Reduktisomerase kann in einem einzigen Schitt detektiert werden, indem die Umwandlung von Glycerinaldehyd-3-phosphat zu 2-C-Methylerythritol-4-phosphat bestimmt wird.

Analog erfolgt die Bestimmung der Aktivitäten von DOXP-Synthase und DOXP-Reduktoisomerase. Für die Bestimmung der DOXP-Synthase-Aktivität eignen sich auch fluorimetrische Verfahren, wie von Querol et al. beschrieben (Querol et al. Abstracts 4th european symposium on plant isoprenoids, Barcelona 21-23 April 1999).

Patentansprüche

- DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 2 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:2, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind.
- 2. DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 4 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:4, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind.
- 3. DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 6 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 6, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind.
- 4. DNA-Sequenz gemäß einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß sie außerdem funktionelle Regulationssignale, insbesondere Promotoren, Operatoren, Enhancer, ribosomale Bindungsstellen, aufweist.
- 5. DNA-Sequenz mit folgenden Teilsequenzen
 - i) Promotor, der in Viren, Eukaryonten und Prokaryonten aktiv ist und die Bildung einer RNA im vorgesehenen Zielgewebe oder den Zielzellen sicherstellt,
 - ii) DNA-Sequenzen gemäß einem der Ansprüche 1 bis 3,
 - iii) 3'-nichttranslatierte Sequenz, die in Viren, Eukaryonten und Prokaryonten zur Addition von Poly-A Resten an das 3'-Ende der RNA führt.
- 6. Verfahren zur Herstellung von transgenen Viren, Eukaryonten und Prokaryonten zur Veränderung des Isoprenoid-Gehaltes, dadurch gekennzeichnet, daß eine DNA-Sequenz gemäß Anspruch 4 oder 5 in das Genom von Viren, eukaryontischen und proka-

WO 00/17233 PCT/EP99/07055 -11-

ryontischen Zellen mit oder ohne Verwendung eines Vektors transferiert und eingebaut wird.

- 7. Transgene Systeme, insbesondere Pflanzen und Pflanzenzellen, welche ein oder mehrere DNA-Sequenzen gemäß der Ansprüche 1 bis 5 als "fremde" oder "zusätzliche" DNA enthalten, die exprimiert werden.
- 8. Expressionsvektor, enthaltend eine oder mehrere DNA-Sequenzen gemäß Anspruch 1 bis 5.
- 9. Protein, welches am 1-Deoxy-D-Xylulose-5-Phosphat-Stoffwechselweges beteiligt ist und a) codiert wird von den DNA-Sequenzen SEQ ID NO: 1,3 oder 5 oder b) codiert wird von DNA-Sequenzen, die mit den DNA-Sequenzen SEQ ID NO: 1,3,5 oder Fragmenten dieser DNA-Sequenzen im DNA-Bereich, der für das reife Protein codiert, hybridisieren.
- 10. Protein nach den Anspruch 9, erhältlich aus den Kulturüberständen von Parasiten oder aus den aufgeschlossenen Parasiten und Aufreinigung über chromatographische und elektrophoretische Techniken.
- 11. Protein nach einem der Ansprüche 9 und 10, dadurch gekennzeichnet, daß es a) das Produkt einer viralen, prokaryontischen oder eukaryontischen Expression einer exogenen DNA ist, b) codiert wird von den Sequenzen SEQ ID NO: 1, 3 oder 5 oder codiert wird von DNA-Sequenzen, die mit den in den DNA-Sequenzen SEQ ID NO: 1, 3 oder 5 oder Fragmenten dieser DNA-Sequenzen im DNA-Bereich, der für das reife Protein kodiert, hybridisieren, oder c) codiert wird von DNA-Sequenzen, die ohne Degeneration des genetischen Codes mit den in b) definierten Sequenzen hybridisieren würden und für ein Polypeptid mit entsprechender Aminosäure-Sequenz kodieren.

- 12. Protein gemäß einem der vorangehenden Ansprüchen, dadurch gekennzeichnet, daß es die Aminosäuresequenzen SEQ ID NO: 2, 4 oder 6 aufweist.
- 13. Verfahren zur Bestimmung der enzymatischen Aktivität des gcpE-Proteins, dadurch gekennzeichnet, daß Phosphorylierung eines Zuckers oder eines Phosphorzuckers oder einer Vorstufe der Isoprenoidbiosynthese, insbesondere die Phosphorylierung von 2-C-Methyl-D-erythritol, 2-C-Methyl-D-erythritol-erythritol-phosphat, insbesondere 2-C-Methyl-D-erythritol-4-phosphat, 2-C-Methyl-D-erythrose, 2-C-Methyl-D-erythrosephosphat, insbesondere 2-C-Methyl-D-erythrose-4-phosphat, und der Phosphat- und Alkoholvorstufen, detektiert wird.
- 14. Verfahren nach Anspruch 13, dadurch gekennzeichnet, daß die Phosphorylierung der folgenden Phosphate oder Alkohole detektiert wird:

```
CH_2(OH) - C(CH_3) = C(OH) - CH_2 - O - PO(OH)_2
CH_{2}(OH) - C(CH_{3}) = C(OH) - CH_{2} - OH
CH<sub>2</sub> (OH) -CH (CH<sub>3</sub>) -CO-CH<sub>2</sub>-O-PO (OH)<sub>2</sub>, CH<sub>2</sub> (OH) -CH (CH<sub>3</sub>) -CO-CH<sub>2</sub>-OH
CH_2=C(CH_3)-CO-CH_2-O-PO(OH)_2, CH_2=C(CH_3)-CO-CH_2-OH,
CH_2=C(CH_3)-CH(OH)-CH_2-O-PO(OH)_2, CH_2=C(CH_3)-CH(OH)-CH_2-OH,
CH_2(OH) - C(=CH_2) - C(OH) - CH_2 - O - PO(OH)_2
CH_{2}(OH) - C (=CH_{2}) - C (OH) - CH_{2} - OH
CHO-CH(CH_3)-CH(OH)-CH_2-O-PO(OH)_2, CHO-CH(CH_3)-CH(OH)-CH_2-OH,
CH_2(OH) - C(OH)(CH_3) - CH = CH - O - PO(OH)_2
CH_2 (OH) -C (OH) (CH<sub>3</sub>) -CH=CH-OH
CH(OH) = C(CH_3) - CH(OH) - CH_2 - O - PO(OH)_2
CH(OH) = C(CH_3) - CH(OH) - CH_2 - OH
(CH_3)_2HC-CO-CH_2-O-PO(OH)_2,
(CH<sub>3</sub>)<sub>2</sub>HC-CO-CH<sub>2</sub>-O-H,
(CH_3)_2HC-CH(OH)-CH_2-O-PO(OH)_2,
(CH_3)_2HC-CH(OH)-CH_2-O-H.
```

15. Verfahren zur gekoppelten Bestimmung der enzymatischen Aktivität der DOXP-Synthase und der DOXP-Reduktase, dadurch gekennzeichnet, daß die Umwandlung von Glycerinaldehyd-3-phosphat zu 2-C-Methylerythritol-4-phosphat detektiert wird.

- 16. Verfahren zum Screening einer Verbindung für die Therapie von infektiösen Prozessen bei Mensch und Tier, wobei das Verfahren umfaßt:
 - a) Bereitstellen einer Wirtszelle, die einen rekombinanten Expressionsvektor enthält, wobei der Vektor zumindest einen Teil der Olignukleotidsequenz gemäß SEQ ID NO:1, SEQ ID NO: 3 oder SEQ ID NO: 5 oder Varianten oder Analoga dieser aufweist, und außerdem eine Verbindung, von der vermutet wird, daß sie eine antimykotische, antibiotische, antiparasitäre oder antivirale Wirkung bei Mensch und Tier hat,
 - b) In-Kontakt-Bringen der Wirtszelle mit der Verbindung und
 - c) Bestimmung der antimikrobiellen, antimykotischen, antibiotischen, antiparasitären oder antiviralen Wirksamkeit der Verbindung.
- 17. Verfahren zum Screening nach Verbindungen zur Behandlung von Pflanzen, wobei das Verfahren umfaßt:
 - a) Bereitstellen einer Wirtszelle, die einen rekombinanten Expressionsvektor enthält, wobei der Vektor zumindest einen Teil der Olignukleotidsequenz gemäß SEQ ID NO:1, SEQ ID NO: 3 oder SEQ ID NO: 5 oder Varianten oder Analoga dieser aufweist, und außerdem eine Verbindung, von der vermutet wird, daß sie eine antimikrobielle, antivirale, antiparasitäre, bakterizide, fungizide oder herbizide Wirkung bei Pflanzen hat,
 - b) In-Kontakt-Bringen der Wirtszelle mit der Verbindung und
 - c) Bestimmung der antimikrobiellen, antiviralen, antiparasitären, bakteriziden, fungiziden oder herbiziden Wirksamkeit der Verbindung.
- 18. Verwendung von DNA nach einem der Ansprüche 1 bis 5 oder von Proteinen nach einem der Ansprüche 9 bis 12 oder von transgenen Systemen nach Ansprüch 7 zur Vorbeugung oder Therapie von Erkrankungen bei Mensch und Tier.

SEQUENZPROTOKOLL

```
<110> Jomaa, Hassan
<120> Gene des 1-Desoxy-D-xylulose-Biosynthesewegs
<130> 15696
<140> PCT/EP99
<141> 1999-09-22
<150> DE19923567.8
<151> 1999-05-22
<150> DE19843279.8
<151> 1998-09-22
<160> 6
<170> PatentIn Ver. 2.1
<210> 1
<211> 1467
<212> DNA
<213> Plasmodium falciparum
<220>
<221> CDS
<222> (1)..(1467)
<220>
<221> gene
<222> (1)..(1467)
<220>
<221> mRNA
<222> (1)..(1467)
<400> 1
Met Lys Lys Tyr Ile Tyr Ile Tyr Phe Phe Ile Thr Ile Thr Ile
 1
                5
                                 10
                                                   15
```

| aat | gat | tta | gta | ata | aat | aat | aca | tca | aaa | tgt | gtt | tcc | att | gaa | aga | 96 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Asp | Leu | Val | Ile | Asn | Asn | Thr | Ser | Lys | Cys | Val | Ser | Ile | Glu | Arg | |
| | | | 20 | | | | | 25 | | | | | 30 | | | |
| • | | | | | | | | | | | | | | | | |
| aga | aaa | aat | aac | gca | tat | ata | aat | tat | ggt | ata | gga | tat | aat | gga | cca | 144 |
| Arg | īуs | Asn | Asn | Ala | Tyr | Ile | Asn | Tyr | Gly | Ile | Gly | Tyr | Asn | Gly | Pro | |
| | | 35 | | | | | 40 | | | | | 45 | | | | |
| | | | | | | | ٠ | | | | | | | | | |
| gat | aat | aaa | ata | aca | aag | agt | aga | aga | tgt | aaa | aga | ata | aag | tta | tgc | 192 |
| Asp | Asn | Lys | Ile | Thr | Lys | Ser | Arg | Arg | Cys | Lys | Arg | Ile | Lys | Leu | Cys | |
| | 50 | | | | | 55 | | | | | 60 | | | | | |
| | | | | | | | | | | | ٠. | | | | | |
| aaa | aag | gat | tta | ata | gat | att | ggt | gca | ata | aag | aaa | cca | att | aat | gta | 240 |
| Lyş | Lys | Asp | Leu | Ile | Asp | Ile | Gly | Ala | Ile | Lys | Lys | Pro | Ile | Asn | Val | |
| 65 | | | | | 70 | • | | | | 75 | | | | | 80 | |
| | | | | | | | | | | | | | | | | |
| gca | att | ttt | gga | agt | act | ggt | agt | ata | ggt | acg | aat | gct | tta | aat | ata | 288 |
| Ala | Ile | Phe | Gly | Ser | Thr | Gly | Ser | Ile | Gly | Thr | Asn | Ala | Leu | Asn | Ile | |
| | | | | 85 | | | | | 90 | | | | | 95 | | |
| | | | | | | | | | | | | | | | | |
| ata | agg | gag | tgt | aat | aaa | att | gaa | aat | gtt | ttt | aat | gtt | aaa | gca | ttg | 336 |
| Ile | Arg | Glu | Cys | Asn | Lys | Ile | Glu | Asn | Val | Phe | Asn | Val | Lys | Ala | Leu | |
| | | | 100 | | | | | 105 | | | | | 110 | | | |
| | | | | | | | | | | | | | | | | |
| tat | gtg | aat | aag | agt | gtg | aat | gaa | tta | tat | gaa | caa | gct | aga | gaa | ttt | 384 |
| Tyr | Val | Asn | Lys | Ser | Val | Asn | Glu | Leu | Tyr | Glu | Gln | Ala | Arg | Glu | Phe | |
| | | 115 | | | | | 120 | | | | | 125 | | | | |
| | • | | | | | | | | | | | | | | | |
| tta | cca | gaa | tat | ttg | tgt | ata | cat | gat | aaa | agt | gta | tat | gaa | gaa | tta | 432 |
| Leu | Pro | Glu | Tyr | Leu | Cys | Ile | His | Asp | Lys | Ser | Val | Tyr | Glu | Glu | Leu | |
| | 130 | | | | | 135 | | | | | 140 | | | | | |
| | | | | | | | | | | | | | | | | |
| aaa | gaa | ctg | gta | aaa | aat | ata | aaa | gat | tat | aaa | cct | ata | ata | ttg | tgt | 480 |
| Lys | Glu | Leu | ۷al | Lys | Asn | Ile | Lys | Asp | Tyr | Lys | Pro | Ile | Ile | Leu | Cys | |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 | |
| | | | | | | | | | | | | | | | | |
| ggt | gat | gaa | ggg | atg | aaa | gaa | ata | tgt | agt | agt | aat | agt | ata | gat | aaa | 528 |
| Gly | Asp | Glu | Gly | Met | Lys | Glu | Ile | Cys | Ser | Ser | Asn | Ser | Ile | Asp | Lys | |
| | | | | 165 | | | | | 170 | | | | | 175 | | |
| | | | | | | | | | | | | | | | | |
| ata | att | att | aat | att | gat | tct | ttt | caa | aaa | tta | tat | tct | act | ato | tat | 576 |

| gca att atg aat aat aat aat aat agt gcg tta gct act aaa aaa gaa tcc att Ala Ile Met Asn Asn Lys Ile Val Ala Leu Ala Asn Lys Glu Ser Ile 200 624 gtc tct gct ggt ttc ttt tta aag aaa tta tta aat att cat aaa aat 672 200 205 205 672 <t< th=""><th>Ile</th><th>Val</th><th>Ile</th><th>Gly 180</th><th>Ile</th><th>Asp</th><th>S r</th><th>Phe</th><th>Gln 185</th><th>Gly</th><th>Leu</th><th>Tyr</th><th>Ser</th><th>Thr 190</th><th></th><th>Tyr</th><th></th></t<> | Ile | Val | Ile | Gly 180 | Ile | Asp | S r | Phe | Gln 185 | Gly | Leu | Tyr | Ser | Thr 190 | | Tyr | |
|---|---|---|--|--|--|--|----------------------------------|--|--|---------------------------------|---|--|---|--|---|--|--------------------|
| 195 200 205 205 207 | gca | att | atg | aat | aat | aaa | ata | gtt | gcg | tta | gct | aat | aaa | gaa | tcc | att | 624 |
| gtc tct gct ggt ttc ttt tta aag aaa tta tta aat att cat aaa aat 672 Val Ser Ala Gly Phe Phe Leu Lys Lys Leu Leu Asn Ile His Lys Asn 210 215 220 720 <td>Ala</td> <td>Ile</td> <td>Met</td> <td>Asn</td> <td>Asn</td> <td>Lys</td> <td>Ile</td> <td>Val</td> <td>Ala</td> <td>Leu</td> <td>Ala</td> <td>Asn</td> <td>Lys</td> <td>Glu</td> <td>Ser</td> <td>Ile</td> <td></td> | Ala | Ile | Met | Asn | Asn | Lys | Ile | Val | Ala | Leu | Ala | Asn | Lys | Glu | Ser | Ile | |
| Val Ser Ala Gly Phe Phe Leu Lys Leu Asn Ile His Lys Asn 210 215 220 225 220 225 226 2270 226 227 220 226 227 220 226 227 220< | | | 195 | | | | | 200 | | | | | 205 | | | | • |
| gca aag ata ata cct gtt gat tca gaa cat agt gct ata ttt caa tgt 720 Ala Lys Ile Ile Pro Val Asp Ser Glu His Ser Ala Ile Phe Gln Cys 235 240 tta gat aat aat aag gta tta aaa aca aaa tgt tta caa gac aat ttt 768 Leu Asp Asn Asn Lys Val Leu Lys Thr Lys Cys Leu Gln Asp Asn Phe 245 250 255 tct aaa att aac aat at aat aat aat aaa at aa aat at ttt tta tgt tca tct gga ggt 816 816 Ser Lys Ile Asn Asn Ile Asn Ile Asn Lys Ile Phe Leu Cys Ser Ser Gly Gly 265 270 cca ttt caa aat tta act at act at ggac gaa tta aaa aat gta aca tca gaa 864 Pro Phe Gln Asn Leu Thr Met Asp Glu Leu Lys Asn Val Thr Ser Glu 275 aat gct tta aag cat cct aaa tgg aaa atg ggt aag aaa at act act ata 285 Asn Ala Leu Lys His Pro Lys Trp Lys Met Gly Lys Lys Ile Thr Ile 290 gat tct gca act atg atg aat aaa ggt tta gag gtt ata gaa acc cat Asp Ser Ala Thr Met Asn Lys Gly Leu Glu Val Ile Glu Thr His 305 305 310 310 315 325 330 335 326 | gtc | tct | gct | ggt | ttc | ttt | tta | aag | aaa | tta | tta | aat | att | cat | aaa | aat | 672 |
| gca aag ata ata cct gtt gat tca gaa cat agt gct ata ttt caa tgt 720 Ala Lys Ile Ile Pro Val Asp Ser Glu His Ser Ala Ile Phe Gln Cys 230 235 240 225 230 235 240 240 tta gat aat aat aat aag gta tta aaa aca aaa tgt tta caa gac aat ttt 768 240 240 tta gat aat aat aat aat aat aat aat aat | Val | Ser | Ala | Gly | Phe | Phe | Leu | Lys | Lys | Leu | Leu | Asn | Ile | His | Lys | Asn | |
| Ala Lys Ile Ile Pro Val Asp Ser Glu His Ser Ala Ile Phe Gln Cys 225 | | 210 | | | | | 215 | | | | | 220 | | | | | |
| 225 | gca | aag | ata | ata | cct | gtt | gat | tca | gaa | cat | agt | gct | ata | ttt | caa | tgt | 720 |
| tta gat aat aat aag gta tta aaa aca aaa tgt tta caa gac aat ttt 768 Leu Asp Asn Asn Lys Val Leu Lys Thr Lys Cys Leu Gln Asp Asn Phe 245 | Ala | Lys | Ile | Ile | Pro | Val | Asp | Ser | Glu | His | Ser | Ala | Ile | Phe | Gln | Cys | |
| Leu Asp Asn Asn Lys Val Leu Lys Thr Lys Cys Leu Gln Asp Asn Phe 245 | 225 | | | | | 230 | | | | | 235 | | | | | 240 | |
| 245 250 255 255 266 267 267 268 268 270 268 270 268 270 268 270 | tta | gat | aat | aat | aag | gta | tta | aaa | aca | aaa | tgt | tta | caa | gac | aat | ttt | 768 |
| CCC aaa at t aac aat at aac aat at aac aat at aaa at at tt tta tgt tca tct gga ggt 816 Ser Lys Ile Asn Asn Ile Asn Lys Ile Phe Leu Cys Ser Ser Gly Gly 260 260 265 270 864 CCa ttt caa aat tta act at gac gac gaa tta aaa aat gta aca tca gaa Phe Gln Asn Leu Thr Met Asp Glu Leu Lys Asn Val Thr Ser Glu 285 285 864 aat gct tta aag cat cct aaa tgg aaa at ggt aag aaa at act at act at get act Lys His Pro Lys Trp Lys Met Gly Lys Lys Ile Thr Ile 290 912 gat tct gca act at g at gat aat aaa ggt tta gag gtt at gaa acc cat Asp Ser Ala Thr Met Met Asn Lys Gly Leu Glu Val Ile Glu Thr His 305 310 315 320 ttt tta ttt gat gta gat tat aat gat at aga gtt at gaa gtt tta ttt ta ttt gat gta gat tat aat gat at gaa gtt at gta gta | Leu | Asp | Asn | Asn | Lys | Val | Leu | Lys | Thr | Lys | Cys | Leu | Gln | Asp | Asn | Phe | |
| Ser Lys 11e Asn Asn Ile Asn Lys 11e Asn Lys 11e Phe Leu Cys Ser Ser Gly Gly 270 cca ttt caa aat tta act atg gac gaa tta aaa aat gta aca tca gaa Pro Phe Gln Asn Leu Thr Met Asp Glu Leu Lys Asn Val Thr Ser Glu 285 864 aat gct tta aag cat cct aaa tgg aaa atg ggt aag aaa atg ggt aag aaa ata act ata 285 285 aat gct tta aag cat cct aaa tgg aaa atg ggt aag aaa atg ggt lys Lys Lys Ile Thr Ile 300 912 Asn Ala Leu Lys His Pro Lys Trp Lys Met Gly Lys Lys Lys Ile Thr Ile 300 960 gat tct gca act atg atg aat aaa ggt tta gag gtt ata gaa acc cat 320 960 Asp Ser Ala Thr Met Met Asn Lys Gly Leu Glu Val Ile Glu Thr His 305 310 305 310 310 4 ttt tta ttt gat gat gat gat tat aat gat g | | | | | 245 | | | | | 250 | | | | | 255 | | |
| Ser Lys 11e Asn Asn Ile Asn Lys 11e Asn Lys 11e Phe Leu Cys Ser Ser Gly Gly 270 cca ttt caa aat tta act atg gac gaa tta aaa aat gta aca tca gaa Pro Phe Gln Asn Leu Thr Met Asp Glu Leu Lys Asn Val Thr Ser Glu 285 864 aat gct tta aag cat cct aaa tgg aaa atg ggt aag aaa atg ggt aag aaa ata act ata 285 285 aat gct tta aag cat cct aaa tgg aaa atg ggt aag aaa atg ggt lys Lys Lys Ile Thr Ile 300 912 Asn Ala Leu Lys His Pro Lys Trp Lys Met Gly Lys Lys Lys Ile Thr Ile 300 960 gat tct gca act atg atg aat aaa ggt tta gag gtt ata gaa acc cat 320 960 Asp Ser Ala Thr Met Met Asn Lys Gly Leu Glu Val Ile Glu Thr His 305 310 305 310 310 4 ttt tta ttt gat gat gat gat tat aat gat g | | | | | | | | | | | • ,• | | | | | | +*,* |
| 260 265 270 266 267 270 267 270 | tct | aaa | att | aac | aat | ata | aat | aaa | ata | ttt | tta | tgt | tca | tct | gga | ggt | 816 💝 |
| CCa ttt Caa aat tta act atg gac gaa tta aaa aat gta aca tca gaa 864 Pro Phe GIn Asn Leu Thr Met Asp GIu Leu Lys Asn Val Thr Ser GIu 285 aat gct tta aag cat cct aaa tgg aaa atg ggt aag aaa ata act ata 912 Asn Ala Leu Lys His Pro Lys Trp Lys Met Gly Lys Lys Ile Thr Ile 290 gat tct gca act atg atg aat aaa ggt tta gag gtt ata gaa acc cat 960 Asp Ser Ala Thr Met Met Asn Lys Gly Leu Glu Val Ile Glu Thr His 305 ttt tta ttt gat gta gat tat aat gat ata gaa gtt ata ga gtt Ash gas Cat aaa 1008 Phe Leu Phe Asp Val Asp Tyr Asn Asp Ile Glu Val Ile Val His Lys 325 gaa tgc att ata cat tct tgt gtt gaa ttt ata gac aaa tca gta ata 1056 Glu Cys Ile Ile His Ser Cys Val Glu Phe Ile Asp Lys Ser Val Ile | Ser | Lys | Ile | Asn | Asn | Ile | Asn | Lys | Ile | Phe | Leu | Суs | Ser | Ser | Gly | бlу | inger i see |
| Pro Phe Gln Asn Leu Thr Met Asp Glu Leu Lys Asn Val Thr Ser Glu aat gct tta aag cat cct aaa tgg aag aag aaa ata act ata 912 Asn Ala Leu Lys His Pro Lys Trp Lys Met Gly Lys Lys Ile Thr Ile Thr Ile Plo gat tct gca act | | | | | | | | | | | | | | | | | |
| aat gct tta aag cat cct aaa tgg aaa atg ggt aag aaa ata act ata 912 Asn Ala Leu Lys His Pro Lys Trp Lys Met Gly Lys Lys Ile Thr Ile 290 gat tct gca act atg atg aat aaa ggt tta gag gtt ata gaa acc cat Asp Ser Ala Thr Met Met Asn Lys Gly Leu Glu Val Ile Glu Thr His 305 ttt tta ttt gat gta gat tat aat gat ata ga ggt tta gag gtt ata gag gt ata tta t | | | | 260 | | | | | 265 | | | | • • | 270 | | | |
| aat gct tta aag cat cct aaa tgg aaa atg ggt aag aaa ata act ata 2912 Asn Ala Leu Lys His Pro Lys Trp Lys Met Gly Lys Lys Ile Thr Ile 290 | cca | ttt | caa | | tta | act | atg | gac | | tta | aaa | aat | gta | | tca | gaa | 864 |
| Ash Ala Leu Lys His Pro Lys Trp Lys Met Gly Lys Lys Ile Thr Ile 290 | | | | aat | | | _ | - | gaa | | | | - | aca | | - | 864 |
| gat tct gca act atg atg aat aaa ggt tta gag gtt ata gaa acc cat 960 Asp Ser Ala Thr Met Met Asn Lys Gly Leu Glu Val Ile Glu Thr His 305 | | | Gln | aat | | | _ | Asp | gaa | | | | Val | aca | | - | 864 |
| gat tot goa act atg atg aat aaa ggt tta gag gtt ata gaa acc cat 960 Asp Ser Ala Thr Met Met Asn Lys Gly Leu Glu Val Ile Glu Thr His 305 310 310 315 320 ttt tta ttt gat gta gat tat aat gat ata gaa gtt ata gta cat aaa 1008 Phe Leu Phe Asp Val Asp Tyr Asn Asp Ile Glu Val Ile Val His Lys 325 330 335 gaa tgc att ata cat tot tgt gtt gaa ttt ata gac aaa toa gta ata 1056 Glu Cys Ile Ile His Ser Cys Val Glu Phe Ile Asp Lys Ser Val Ile | Pro | Phe | G1n 275 | aat Asn | Leu | Thr | Met | Asp 280 | gaa Glu | Leu | Lys | Asn | Val 285 | aca Thr | Ser | Glu | |
| Asp Ser Ala Thr Met Met Asn Lys Gly Leu Glu Val Ile Glu Thr His 305 310 310 315 320 ttt tta ttt gat gta gat tat aat gat ata gaa gtt ata gta cat aaa 1008 Phe Leu Phe Asp Val Asp Tyr Asn Asp Ile Glu Val Ile Val His Lys 325 330 335 gaa tgc att ata cat tct tgt gtt gaa ttt ata gac aaa tca gta ata 1056 Glu Cys Ile Ile His Ser Cys Val Glu Phe Ile Asp Lys Ser Val Ile | Pro | Phe gct | Gln 275 tta | aat Asn aag | Leu cat | Thr cct | Met | Asp 280 tgg | gaa Glu aaa | Leu atg | Lys ggt | Asn aag | Val 285 | aca Thr | Ser | Glu | |
| 310 315 320 ttt tta ttt gat gta gat tat aat gat ata gaa gtt ata gta cat aaa 1008 Phe Leu Phe Asp Val Asp Tyr Asn Asp Ile Glu Val Ile Val His Lys 325 330 335 gaa tgc att ata cat tct tgt gtt gaa ttt ata gac aaa tca gta ata 1056 Glu Cys Ile Ile His Ser Cys Val Glu Phe Ile Asp Lys Ser Val Ile | Pro | Phe gct Ala | Gln 275 tta | aat Asn aag | Leu cat | Thr cct | Met aaa Lys | Asp 280 tgg | gaa Glu aaa | Leu atg | Lys ggt | Asn aag Lys | Val 285 | aca Thr | Ser | Glu | |
| ttt tta ttt gat gta gat tat aat gat ata gaa gtt ata gta cat aaa 1008 Phe Leu Phe Asp Val Asp Tyr Asn Asp Ile Glu Val Ile Val His Lys 325 330 335 gaa tgc att ata cat tct tgt gtt gaa ttt ata gac aaa tca gta ata 1056 Glu Cys Ile Ile His Ser Cys Val Glu Phe Ile Asp Lys Ser Val Ile | Pro aat Asn | gct Ala 290 | Gln 275 tta Leu | aat Asn aag Lys | Leu cat His | Thr cct Pro | Met aaa Lys 295 | Asp 280 tgg Trp | gaa Glu aaa Lys | Leu atg Met | Lys ggt Gly | aag Lys 300 | Val 285 aaa Lys | aca Thr ata Ile | Ser act Thr | Glu ata Ile | 912 |
| Phe Leu Phe Asp Val Asp Tyr Asn Asp Ile Glu Val Ile Val His Lys 325 330 335 gaa tgc att ata cat tct tgt gtt gaa ttt ata gac aaa tca gta ata 1056 Glu Cys Ile Ile His Ser Cys Val Glu Phe Ile Asp Lys Ser Val Ile | Pro aat Asn gat | gct Ala 290 | Gln 275 tta Leu gca | aat Asn aag Lys | Leu cat His | Thr cct Pro | Met aaa Lys 295 | Asp 280 tgg Trp | gaa Glu aaa Lys | Leu atg Met | Lys ggt Gly gag | aag Lys 300 gtt | Val 285 aaa Lys | aca Thr ata Ile | Ser act Thr | Glu ata Ile | 912 |
| Phe Leu Phe Asp Val Asp Tyr Asn Asp Ile Glu Val Ile Val His Lys 325 330 335 gaa tgc att ata cat tct tgt gtt gaa ttt ata gac aaa tca gta ata 1056 Glu Cys Ile Ile His Ser Cys Val Glu Phe Ile Asp Lys Ser Val Ile | Pro aat Asn gat Asp | gct Ala 290 | Gln 275 tta Leu gca | aat Asn aag Lys | Leu cat His | Thr cct Pro atg Met | Met aaa Lys 295 | Asp 280 tgg Trp | gaa Glu aaa Lys | Leu atg Met | Lys ggt Gly gag Glu | aag Lys 300 gtt | Val 285 aaa Lys | aca Thr ata Ile | Ser act Thr | Glu ata Ile cat | 912 |
| 325 330 335 gaa tgc att ata cat tct tgt gtt gaa ttt ata gac aaa tca gta ata 1056 Glu Cys Ile Ile His Ser Cys Val Glu Phe Ile Asp Lys Ser Val Ile | Pro aat Asn gat Asp | gct Ala 290 | Gln 275 tta Leu gca | aat Asn aag Lys | Leu cat His | Thr cct Pro atg Met | Met aaa Lys 295 | Asp 280 tgg Trp | gaa Glu aaa Lys | Leu atg Met | Lys ggt Gly gag Glu | aag Lys 300 gtt | Val 285 aaa Lys | aca Thr ata Ile | Ser act Thr | Glu ata Ile cat | 912 |
| gaa tgc att ata cat tct tgt gtt gaa ttt ata gac aaa tca gta ata 1056 Glu Cys Ile Ile His Ser Cys Val Glu Phe Ile Asp Lys Ser Val Ile | Pro aat Asn gat Asp 305 | gct Ala 290 tct Ser | Gln 275 tta Leu gca Ala | aat Asn aag Lys act Thr | Cat His atg Met | Thr cct Pro atg Met 310 | Met aaa Lys 295 aat Asn | Asp 280 tgg Trp aaa Lys | gaa Glu aaa Lys ggt Gly | Leu atg Met tta Leu | ggt Gly gag Glu 315 | aag Lys 300 gtt Val | Val 285 aaa Lys ata Ile | aca Thr ata Ile gaa Glu | act Thr | Glu ata Ile cat His 320 | 912 960 |
| Glu Cys Ile Ile His Ser Cys Val Glu Phe Ile Asp Lys Ser Val Ile | aat Asn gat Asp 305 | gct Ala 290 tct Ser | Gln 275 tta Leu gca Ala | aat Asn aag Lys act Thr | Cat His atg Met | Thr cct Pro atg Met 310 gat | Met aaa Lys 295 aat Asn | Asp 280 tgg Trp aaa Lys | gaa Glu aaa Lys ggt Gly | Leu atg Met tta Leu | ggt Gly gag Glu 315 | aag Lys 300 gtt Val | Val 285 aaa Lys ata Ile | aca Thr ata Ile gaa Glu | act Thr acc Thr | Glu ata Ile cat His 320 | 912 960 |
| | aat Asn gat Asp 305 | gct Ala 290 tct Ser | Gln 275 tta Leu gca Ala | aat Asn aag Lys act Thr | Cat His atg Met | Thr cct Pro atg Met 310 gat | Met aaa Lys 295 aat Asn | Asp 280 tgg Trp aaa Lys | gaa Glu aaa Lys ggt Gly | Leu atg Met tta Leu ata Ile | ggt Gly gag Glu 315 | aag Lys 300 gtt Val | Val 285 aaa Lys ata Ile | aca Thr ata Ile gaa Glu | Ser act Thr acc Thr | Glu ata Ile cat His 320 | 912 960 |
| 340 345 350 | Pro aat Asn gat Asp 305 ttt | gct Ala 290 tct Ser tta Leu | Gln 275 tta Leu gca Ala ttt | aat Asn aag Lys act Thr | Cat His atg Met Val 325 | Thr cct Pro atg Met 310 gat Asp | Met aaa Lys 295 aat Asn tat | Asp 280 tgg Trp aaa Lys aat Asn | gaa Glu aaa Lys ggt Gly gat Asp | Leu atg Met tta Leu ata Ile 330 | ggt Gly gag Glu 315 gaa Glu | aag Lys 300 gtt Val gtt | Val 285 aaa Lys ata Ile | aca Thr ata Ile gaa Glu gta Val | act Thr acc Thr cat His 335 | Glu ata Ile cat His 320 aaa Lys | 912 960 1008 |
| | aat Asn gat Asp 305 ttt Phe | gct Ala 290 tct Ser tta Leu | Gln 275 tta Leu gca Ala ttt Phe | aat Asn aag Lys act Thr gat Asp | cat His atg Met gta Val 325 | Thr cct Pro atg Met 310 gat Asp | Met aaa Lys 295 aat Asn tat Tyr | Asp 280 tgg Trp aaa Lys aat Asn | gaa Glu aaa Lys ggt Gly gat Asp | Leu atg Met tta Leu ata Ile 330 | ggt Gly gag Glu 315 gaa Glu | aag Lys 300 gtt Val gtt | Val 285 aaa Lys ata Ile ata | aca Thr ata Ile gaa Glu gta Val | act Thr acc Thr cat His 335 | Glu ata Ile cat His 320 aaa Lys | 912 960 1008 |

| - | | | | | | | | | tta Leu 365 | | | | 1104 |
|---|---|---|-------------------|---|---|-----|---|-----|-------------------|---|---|-------------------|------|
| | | - | - | | | | | | tta Leu | _ | _ | _ | 1152 |
| - | - | | | | | | | Leu | gaa Glu | | | <u>.</u> | 1200 |
| - | | | - | | - | | | | aac Asn | | | | 1248 |
| | - | | | | • | | • | | tta Leu | | _ | | 1296 |
| | | | | - | | | | | tcg Ser 445 | | - | | 1344 |
| - | | | | - | - | _ | • | • | gaa Glu | - | | _ | 1392 |
| | | | | | | | | | aaa Lys | - | | gat Asp 480 | 1440 |
| | | | cat His 485 | | | tag | | | | | | | 1467 |

<210> 2 <211> 488

<212> PRT

<213> Plasmodium falciparum

| | | | | | | | | | 5 | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| <400 | > 2 | | | | | | | | | * | | | | • | |
| Met 1 | Lys | Lys | Tyr | Ile 5 | Tyr | Ile | Tyr | Phe | Phe 10 | Phe | Ile | Thr | Ile | Thr 15 | Ile |
| Asn | Asp | Leu | Val 20 | Ile | Asn | Asn | Thr | Ser 25 | Lys | Cys | Val | Ser | 11e 30 | Glu | Arg |
| Arg | Lys | Asn 35 | Asn | Ala | Tyr | Ile | Asn 40 | Tyr | Gly | Ile | Gly | Tyr 45 | Asn | Gly | Pro |
| Asp | Asn 50 | Lys | Ile | Thr | Lys | Ser 55 | Arg | Arg | Cys | Lys | Arg 60 | Ile | Lys | Leu | Cys |
| Lys 65 | Lys | Asp | Leu | Ile | Asp 70 | Ile | Gly | Ala | Ile | Lys 75 | Lys | Pro | Ile | Asn | Val 80 |
| Ala | Ile | Phe | Gly | Ser 85 | Thr | Gly | Ser | Ile | 90 | Thr | Asn | Ala | Leu | Asn 95 | Ile |
| Ile | Arg | Glu | Cys 100 | Asn | Lys | Ile | Glu | Asn 105 | Val | Phe | Asn | Val | Lys 110 | Ala | Leu |
| Tyr | Val | Asn 115 | Lys | Ser | Val | Asn | Glu 120 | Leu | Tyr | Glu | Gln | Ala 125 | Arg | Glu | Phe |
| Leu | Pro 130 | Glu | Tyr | Leu | Cys | Ile 135 | His | Asp | Lys | Ser | Val 140 | Tyr | Glu | Glu | Leu |
| Lys 145 | Glu | Leu | Val | Lys | Asn 150 | Ile | Lys | Asp | Tyr | Lys 155 | Pro | Ile | Ile | Leu | Cys 160 |
| Gly | Asp | Glu | Gly | Met 165 | Lys | Glu | Ile | Cys | Ser 170 | Ser | Asn | Ser | Ile | Asp 175 | Lys |
| Ile | Val | Ile | Gly 180 | Ile | Asp | Ser | Phe | Gln 185 | Gly | Leu | Tyr | Ser | Thr 190 | Met | Tyr |

Ala Ile Met Asn Asn Lys Ile Val Ala Leu Ala Asn Lys Glu Ser Ile 200

Val Ser Ala Gly Phe Phe Leu Lys Lys Leu Leu Asn Ile His Lys Asn

205

210 215 220

Ala Lys Ile Ile Pro Val Asp Ser Glu His Ser Ala Ile Phe Gln Cys 225 230 235 240

Leu Asp Asn Asn Lys Val Leu Lys Thr Lys Cys Leu Gln Asp Asn Phe
245 250 255

Ser Lys Ile Asn Asn Ile Asn Lys Ile Phe Leu Cys Ser Ser Gly Gly
260 265 270

Pro Phe Gln Asn Leu Thr Met Asp Glu Leu Lys Asn Val Thr Ser Glu 275 280 285

Asn Ala Leu Lys His Pro Lys Trp Lys Met Gly Lys Lys Ile Thr Ile 290 295 300

Asp Ser Ala Thr Met Met Asn Lys Gly Leu Glu Val Ile Glu Thr His 305 310 315 320

Phe Leu Phe Asp Val Asp Tyr Asn Asp Ile Glu Val Ile Val His Lys 325 330 335

Glu Cys Ile Ile His Ser Cys Val Glu Phe Ile Asp Lys Ser Val Ile 340 345 350

Ser Gln Met Tyr Tyr Pro Asp Met Gln Ile Pro Ile Leu Tyr Ser Leu 355 360 365

Thr Trp Pro Asp Arg Ile Lys Thr Asn Leu Lys Pro Leu Asp Leu Ala 370 375 380

Gln Val Ser Thr Leu Thr Phe His Lys Pro Ser Leu Glu His Phe Pro 385 390 395 400

Cys Ile Lys Leu Ala Tyr Gln Ala Gly Ile Lys Gly Asn Phe Tyr Pro 405 410 415

Thr Val Leu Asn Ala Ser Asn Glu Ile Ala Asn Asn Leu Phe Leu Asn 420 425 430

Asn Lys Ile Lys Tyr Phe Asp Ile Ser Ser Ile Ile Ser Gln Val Leu

435

445

Glu Ser Phe Asn Ser Gln Lys Val Ser Glu Asn Ser Glu Asp Leu Met 450 455 460

440

Lys Gln Ile Leu Gln Ile His Ser Trp Ala Lys Asp Lys Ala Thr Asp
465 470 475 480

Ile Tyr Asn Lys His Asn Ser Ser 485

<210> 3

<211> 3872

<212> DNA

<213> Plasmodium falciparum

<220>

<221> CDS

<222> (126)..(3740)

<220>

<221> gene

<222> (1)..(3870)

<220>

<221> mRNA

<222> (1)..(3870)

<400> 3

ggtaatatac gtataatata tatataatat attcttacgt atgtatcatt tatgaatcat 60

aataatattc taaatttacc ttccgttttt gctcgatctt ctcattttcg tttcagcttt 120

tatca atg att ttt aat tat gtg ttt ttt aag aac ttt gta cca gtt gtt 170

Met Ile Phe Asn Tyr Val Phe Phe Lys Asn Phe Val Pro Val Val

1 5 10 15

cta tac att ctc ctt ata ata tat att aac tta aat ggc atg aat aat 218 Leu Tyr Ile Leu Leu Ile Ile Tyr Ile Asn Leu Asn Gly Met Asn Asn

20

25

| aaa | aat | caa | ata | aaa | aca | gaa | aaa | att | tat | ata | aag | aaa | ttg | aat | agg | 266 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|-----|
| Lys | Asn | Gln | Ile | Lys | Thr | Glu | Lys | Ile | Tyr | Ile | Lys | Lys | Leu | Asn | Arg | |
| | | | 35 | | | | | 40 | | | | , | 45 | | | |
| , | | | | | | | | | | | | | | | | |
| ttg | tca | agg | aaa | aat | tcg | tta | tgt | agt | tct | aaa | aat | aaa | ata | gca | tgc | 314 |
| Leu | Ser | Arg | Lys | Asn | Ser | Leu | Cys | Ser | Ser | Lys | Asn | Lys | Ile | Ala | Cys | |
| | | 50 | | | | | 55 | | | | | 60 | | | | |
| | | | | | | | | | | | | | | | | |
| ttg | ttc | gat | ata | gga | aat | gat | gat | aat | aga | aat | acg | aca | tat | ggc | tat | 362 |
| Leu | Phe | Asp | Ile | Gly | Asn | Asp | Asp | Asn | Arg | Asn | Thr | Thr | Tyr | Gly | Tyr | |
| | 65 | | | | | 70 | | | | | . 75 | | | | | |
| | | | | | | | | | | | • | | | | | |
| aat | gtg | aat | gtt | aaa | aat | gat | gat | att | aat | tcc | tta | cta | aaa | aat | aat | 410 |
| Asn | Val | Asn | Val | Lys | Asn | Asp | Asp | Ile | Asn | Ser | Leu | Leu | Lys | Asn | Asn | |
| 80 | | | | | 85 | | | | | 90 | | | | | 95 | |
| | | | | | | | | | | | | | | | | |
| tat | agt | aat | aaa | ttg | tac | atg | gat | aag | agg | aaa | aat | att | aat | aat | gta | 458 |
| Tyr | Ser | Asn | Lys | Leu | Tyr | Met | Asp | Lys | Arg | Lys | Asn | Ile | Asn | Asn | Val | |
| | | | | 100 | | | | | 105 | | | | | 110 | | |
| | | | | | | | | | | | | | | | | |
| att | agt | act | aat | aaa | ata | tct | ggg | tcc | att | tca | aat | att | tgt | agt | aga | 506 |
| Ile | Ser | Thr | Asn | Lys | Ile | Ser | Gly | Ser | Ile | Ser | Asn | Ile | Cys | Ser | Arg | |
| | | | 115 | | | | | 120 | | | | | 125 | | | |
| | | | | | | | | | | | | | | | | |
| aat | caa | aaa | gaa | aat | gaa | caa | aaa | aga | aat | aaa | caa | aga | tgt | tta | act | 554 |
| Asn | Gln | Lys | Glu | Asn | Glu | Gln | Lys | Arg | Asn | Lys | Gln | Arg | Cys | Leu | Thr | |
| | | 130 | | | | | 135 | | | | | 140 | | | | |
| | • | | | | | | | | | | | | | | | |
| caa | tgt | cac | act | tat | aat | atg | tca | cat | gaa | cag | gac | aaa | cta | gct | aat | 602 |
| Gln | Cys | His | Thr | Tyr | Asn | Met | Ser | His | Glu | Gln | Asp | Lys | Leu | Ala | Asn | |
| | 145 | | | | | 150 | | | | | 155 | | | | | |
| | | | | | | | | | | | | | | | | |
| gat | aat | aat | agg | aat | aat | aaa | aag | aat | ttt | aat | tta | tta | ttt | ata | aat | 650 |
| Asp | Asn | Asn | Arg | Asn | Asn | Lys | Lys | Asn | Phe | Asn | Leu | Leu | Phe | Ile | Asn | |
| 160 | | | | | 165 | | | | | 170 | | | | | 175 | |
| | | | | | | | | | | | | | | | | |
| tat | ttt | aat | ttg | aaa | cga | atg | aaa | aat | tct | ctt | cta | aat | aaa | gac | aat | 698 |
| Tyr | Phe | Asn | Leu | Lys | Arg | Met | Lys | Asn | Ser | Leu | Leu | Asn | Lys | Asp | Asn | |
| | | | | 180 | | | | | 185 | | | | | 190 | | |
| | | | | | | | | | | | | | | | | |
| ttc | ttt | tac | tgt | aaa | gaa | aaa | aaa | ttg | tca | ttt | ctg | cat | aag | gcc | tat | 746 |

| | | | | | | | | | | , | | | | | | |
|--------------|--------------|-------|-----|-------|--------|-------|------|-------------|-----|-----|------------|-------|-----|-----|--------------|------|
| Phe | Phe | Tyr | Cys | Lys | Glu | Lys | Lys | Leu | Ser | Phe | Leu | His | Lys | Ala | Tyr | |
| | | | 195 | | | | | 200 | | | | | 205 | | | |
| | | | | | | | | | | | | | | | | |
| | 222 | | + | + ~ ~ | act | +++ | C22 | 221 | tat | agt | ++= | 222 | 202 | 222 | tet | 794 |
| | | | | - | | | | | | _ | | | - | | | 7 23 |
| Lys | Lys | - | Asn | Cys | Thr | Phe | | Asn | Tyr | Ser | Leu | _ | Arg | Lys | Ser | |
| | | 210 | | | | | 215 | | | | | 220 | | | | |
| | | | | | | • | | | | | | | | | | |
| aat | cgt | gat | tca | cat | aaa | ttg | ttt | tct | gga | gaa | ttt | gac | gat | tat | aca | 842 |
| Asn | Arg | Asp | Ser | His | Lys | Leu | Phe | Ser | Gly | Glu | Phe | Asp | Asp | Tyr | Thr | |
| | 225 | | | | | 230 | | | | | 235 | | | | | |
| | | | | | | | | | | | | | | | | |
| - | - | 22+ | act | ++= | tat | ~ ~ ~ | tcc | ~ 22 | 222 | | 422 | tac | att | 202 | cta | 890 |
| | | | - | | | • | | • | | | - | | | | | 000 |
| | Asn | Asn | Ala | Leu | Tyr | GIU | ser | GIU | гÀ2 | | GIU | Tyr | ııe | Thr | | |
| 240 | | | | | 245 | | | | | 250 | | | | | 255 | |
| | | | | | | | | | | | | | | | | |
| aat | aat | aat | aat | aaa | aat | aat | aat | aat | aaa | aat | aat | gat | aat | aaa | aat | 938 |
| Asn | Asn | Asn | Asn | Lys | Asn | Asn | Asn | Asn | Lys | Asn | Asn | Asp | Asn | Lys | Asn | |
| | | | | 260 | | | | | 265 | | | | | 270 | | |
| | | | | | | | | | | | | | | | | |
| aat | gat | aat | aat | gat | tat | aat | aat | aat | aat | agt | tat | aat | aat | tta | gga | 986 |
| | - | | | - | Tyr | | | | | - | - | | | | | • |
| ASII | nsp | ASII | | Asp | ıyı | ASII | ASII | | non | Jer | Cys | non | | Dea | Oly | |
| | | | 275 | | | | | 280 | | | | | 285 | | | |
| | | | | | | | | | | | | | | | | |
| gag | aga | tcc | aat | cat | tat | gat | aat | tat | ggt | gga | gat | aat | aat | aat | cca | 1034 |
| Glu | Arg | Ser | Asn | His | Tyr | Asp | Asn | Tyr | Gly | Gly | Asp | Asn | Asn | Asn | Pro | |
| | | 290 | | | | | 295 | | | | | 300 | | | | |
| | | | | | | | | | | | | | | | | |
| tqt | aat | aat | aat | aat | gac | aaa | tat | gat | ata | gga | aaa | tat | ttc | aaa | cag | 1082 |
| | | | | | Asp | | | | | | | | | | | |
| 0,0 | 305 | | | | · ··op | 310 | - 7 | | | 01, | 315 | - 1 - | | -,- | •= | |
| | 303 | | | | | 310 | | | | | 213 | | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | aat | | | | | | | | | | | 1130 |
| Ile | Asn | Thr | Phe | Ile | Asn | Ile | Asp | Glu | Tyr | Lys | Thr | Ile | Tyr | Gly | Asp | |
| 320 | | | | | 325 | | | | | 330 | | | | | 335 | |
| | | | | | | | | | | | | | | | | |
| gaa | ata | tat | aaa | gaa | ata | tat | gaa | cta | tat | gta | gaa | aga | aat | att | cct | 1178 |
| • | | | | - | Ile | | - | | | - | • | - | | | | |
| | | - , - | -,, | 340 | | -] - | | | 345 | | | | | 350 | - | |
| | | | | 740 | | | | | 247 | | | | | JJ0 | | |
| | | | | | | | | | | _ | | | | | | 1000 |
| | | | | | aaa | | | | | | | | | | | 1226 |
| Glu | Tyr | Tyr | Glu | Arg | Lys | Tyr | Phe | Ser | Glu | Asp | Ile | Lys | Lys | Ser | Val | |
| | | | 355 | | | | | 360 | | | | | 365 | | | • |
| | | | | | | | | | | | | | | | | |

| | | | | | | | | | | 10 | | | | | | |
|-----|-----|------|-----|-------|------|-----|-----|------|-----|-------|-----|-----|------|-----------|-----|------|
| cta | ttt | gat | ata | gat | aaa | tat | aat | gat | gtc | gaa | ttt | gaa | aaa | gct | ata | 1274 |
| Leu | Phe | Asp | Ile | Asp | Lys | Tyr | Asn | Asp | Val | Glu | Phe | Glu | Lys | Ala | lle | |
| | | 370 | | | | | 375 | , | | • | | 380 | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | - 4 | | | 1200 |
| | • | - | ttt | | | | | - | | | | | | • | | 1322 |
| Lys | Glu | Glu | Phe | Ile | Asn | Asn | Gly | Val | Tyr | Ile | Asn | Asn | Ile | Asp | Asn | |
| | 385 | | | | | 390 | | | | | 395 | | | | | |
| | | | | | | | | | | | | | | | | |
| aca | tat | tat | aaa | aaa | gaa | aat | att | tta | ata | atg | aaa | aag | ata | tta | cat | 1370 |
| Thr | Tvr | Tvr | Lys | Lvs | Glu | Asn | Ile | Leu | Ile | Met | Lvs | Lvs | Ile | Leu | His | |
| 400 | 4 - | - 4 | -1- | | 405 | | | | | 410 | _ | | | | 415 | |
| 400 | | | | | 405 | | | | | 410 | | | | | 413 | |
| | | | | | | | | | | | | | | | | |
| tat | ttc | cca | tta | tta | aaa | tta | att | aat | aat | cca | tca | gat | tta | aaa | aag | 1418 |
| Tyr | Phe | Pro | Leu | Leu | Lys | Leu | Ile | Asn | Asn | Pro | Ser | Asp | Leu | Lys | Lys | |
| | | | | 420 | | | | | 425 | | | | | 430 | | |
| | | | | | | | | | | | | | | | | |
| tta | aaa | aaa | caa | tat | tta | cct | tta | tta | gca | cat | gaa | tta | aaa | ata | ttt | 1466 |
| | | | Gln | | | | | | | | | | | | | |
| 200 | 2,0 | טעם | | - y - | Dea | | | 440 | | | 014 | 200 | 445 | | | |
| | | | 435 | | | | | 440 | | | | | 443 | | | |
| | | | | | | | | | | | | | | | | |
| tta | ttt | ttt | att | gta | aat | ata | aca | gga | ggt | cat | ttt | tcc | tct | gtt | tta | 1514 |
| Leu | Phe | Phe | Ile | Val | Asn | Ile | Thr | Gly | Gly | His | Phe | Ser | Ser | Val | Leu | |
| | | 450 | | | | | 455 | | | | | 460 | | | | |
| | | | | | | | | | | | | | | | | |
| agc | tct | tta | gaa | att | caa | tta | tta | tta | tta | tat | att | ttt | aat | caa | cca | 1562 |
| | | | Glu | | | | | | | | | | | | | |
| 001 | | Dou | Ozu | 116 | 0111 | | 200 | 200 | 200 | - , - | | | | 02 | | |
| | 465 | | | | | 470 | | | | | 475 | | | | | |
| | | | | | | | | | | | | | | | | |
| tat | gat | aat | gtt | ata | tat | gat | ata | gga | cat | caa | gca | tat | gta | cat | aag | 1610 |
| Tyr | Asp | Asn | Val | Ile | Tyr | Asp | Ile | Gly | His | Gln | Ala | Tyr | Val | His | Lys | • |
| 480 | | | | | 485 | | | | | 490 | | | | | 495 | |
| | | | | | | | | | | | | | | | | |
| ata | tta | acc | gga | 202 | 222 | cta | tta | ttt | cta | tca | tta | aga | aat | aaa | aaa | 1658 |
| | | | Gly | - | | | | | | | | | | | | |
| 116 | Leu | 1111 | GTÀ | _ | гуз | Leu | Leu | FIIC | | Ser | Leu | Arg | non | | цуз | |
| | | | | 500 | | | | | 505 | | | | | 510 | | |
| | | | | | | | | | | | | | | | | |
| ggt | att | agt | gga | ttc | cta | aat | att | ttt | gaa | agt | att | tat | gat | aaa | ttt | 1706 |
| Gly | Ile | Ser | Gly | Phe | Leu | Asn | Ile | Phe | Glu | Ser | Ile | Tyr | Asp | Lys | Phe | |
| | | | 515 | | | | | 520 | | | | | 525 | | | |
| | | | | | | | | | | | | | | | | |
| ~~~ | a | ~~- | | | + | 20+ | +07 | ++- | a~+ | ~~+ | 2+3 | Can | uu s | +=+ | tat | 1754 |
| ygg | gct | ggt | cac | agt | tcc | act | LUd | LLd | ayı | gct | aca | cad | yya | Lat | cac | 1,74 |

| | | | | | | | | | 1 | . 1 | | | | | | |
|-------|------|------------------|------|-------------|-----|-----|-----|-----|-----|-----|------|-----|---------|------|-------|------|
| Gly | Ala | Gly | His | Ser | Ser | Thr | Ser | Leu | Ser | Ala | Ile | Gln | Gly | Tyr | Tyr | |
| | | 530 | | | | | 535 | | | | | 540 | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | 1000 |
| _ | | • - | | | gtg | | | | Ť., | | | | | | _ | 1802 |
| Glu | Ala | Glu | Trp | Gln | Val | Lys | Asn | Lys | Glu | Lys | Tyr | Gly | Asn | Gly | Asp | |
| | 545 | | | | | 550 | | | | | 555 | | | | | |
| | | | | | | | | | | | | | | | | |
| ata | gaa | ata | agt | gat | aac | gca | aat | gtc | acg | aat | aat | gaa | agg | ata | ttt | 1850 |
| Ile | Glu | Ile | Ser | Asp | Asn | Ala | Asn | Val | Thr | Asn | Asn | Glu | Ara | Ile | Phe | |
| 560 | | | | | 565 | | | | | 570 | | | - 10- 9 | | 575 | |
| 300 | | | | | 203 | | | | | 370 | | | | | 373 | |
| | | | | | | | | | | | | | | | | |
| caa | aaa | gga | ata | cac | aat | gat | aat | aat | att | aac | aat | aat | att | aat | aat | 1898 |
| Gln | Lys | Gly | Ile | His | Asn | Asp | Asn | Asn | Ile | Asn | Asn | Asn | Ile | Asn | Asn | |
| | | | | 580 | | | | | 585 | | | | | 590 | | |
| | | | | | | | | | | | | | | | | |
| aat | aat | tat | atc | aat | cct | tca | gat | ata | σta | aga | aga | gaa | aat | aca | aat | 1946 |
| | | | | | Pro | | | | | | | - | | | | |
| 71311 | AJII | | | ASII | 110 | 561 | nop | | *** | Gry | n. g | Gru | | 1111 | ASII | |
| | | | 595 | | | | | 600 | | | | | 605 | | | |
| | | | | | | | | | | | | | | | | |
| gta | cca | aat | gta | cga | aat | gat | aac | cat | aac | gtg | gat | aaa | gta | cac | att | 1994 |
| Val | Pro | Asn | Val | Arg | Asn | Asp | Asn | His | Asn | Val | Asp | Lys | Val | His | Ile | |
| | | 610 | | | | | 615 | | | | | 620 | | | | |
| | | | | | | | | | | | | | | | | |
| qct | att | ata | qqa | gat | ggt | gat | tta | aca | ggt | gga | atq | gca | tta | gaa | gcg | 2042 |
| _ | | | | • | Gly | | | | | | - | - | | _ | | |
| 1114 | 625 | | Ory | nsp | Cry | | 200 | | 017 | 01, | 635 | | Dea | 914 | | |
| | 623 | | | | | 630 | | | | | 635 | | | | | |
| | | | | | | | | | | | | | | | | |
| tta | aat | tat | att | tca | ttc | ttg | aat | tct | aaa | att | tta | att | att | tat | aat | 2090 |
| Leu | Asn | Tyr | Ile | Ser | Phe | Leu | Asn | Ser | Lys | Ile | Leu | Ile | Ile | Tyr | Asn | |
| 640 | | | | | 645 | | | | | 650 | | | | | 655 | |
| | | | | | | | | | | | | | | | | ٠ |
| gat | aac | gga | caa | att | tct | tta | cca | aca | aat | gcc | qta | agt | ata | tca | ggt | 2138 |
| | | | | - | Ser | | | | | | | | | | | |
| | | O ₊ y | 0111 | | 001 | | | | | ••• | , 41 | 501 | | | 027 | |
| | | | | 660 | | | | | 665 | | | | | 670 | | |
| | | | | | | | | | | | | | | | | |
| aat | aga | cct | ata | ggt | tct | ata | tca | gat | cat | tta | cat | tat | ttt | gtt | tct | 2186 |
| Asn | Arg | Pro | Ile | Gly | Ser | Ile | Ser | Asp | His | Leu | His | Tyr | Phe | Val | Ser ´ | |
| | | | 675 | | | | | 680 | | | | | 685 | | | |
| | | | | | | | | | | | | | | | | |
| aa+ | at a | na a | ac > | - - - | gct | aat | aat | aat | 222 | tta | tea | 222 | aat | aca | 222 | 2234 |
| | | - | • | | - | | - | | | | • | | | - | | |
| Asn | тте | | ATa | Asn | Ala | GTÀ | | ASN | гÀ2 | ьeu | ser | | ASN | мта | пÀ2 | |
| | | 690 | | | | | 695 | | | | | 700 | | | | • |

| | | | | | | | | | _ | | | | | | | |
|-------|-----|-------|------|-----|-----|------|-----|-----|----------|-----|----------|-------------|-----|-----|-----|------|
| gag | aat | aac | att | ttt | gaa | aat | ttg | aat | tat | gat | tat | att | ggt | gtt | gtg | 2282 |
| Glu | Asn | Asn | Ile | Phe | Glu | Asn | Leu | Asn | Tyr | Asp | Tyr | Ile | Gly | Val | Val | |
| | 705 | | | | | 710 | | | | | 715 | | | | | |
| | | | | | | | | | | | | | | | | |
| aat | ggt | aat | aat | aca | gaa | gag | ctc | ttt | aaa | gta | tta | aat | aat | ata | aaa | 2330 |
| Asn | Gly | Asn | Asn | Thr | Glu | Glu | Leu | Phe | Lys | Val | Leu | Asn | Asn | Ile | Lys | |
| 720 | | | | | 725 | | | | | 730 | | | | | 735 | |
| | | | | | | | | | | | | | | | | |
| gaa | aat | aaa | tta | aaa | aga | gct | act | gtt | ctt | cat | gta | cgt | aca | aaa | aaa | 2378 |
| Glu | Asn | Lys | Leu | Lys | Arg | Ala | Thr | Val | Leu | His | Val | Arg | Thr | Lys | Lys | |
| | | | | 740 | | | | | 745 | | | | | 750 | | |
| | | | | | | | | | | | | | | | | |
| tcg | aat | gat | ttt | ata | aat | tca | aag | agt | cca | ata | agt | ata | ttg | cac | tct | 2426 |
| Ser | Asn | Asp | Phe | Ile | Asn | Ser | Lys | Ser | Pro | Ile | Ser | Ile | Leu | His | Ser | |
| | | | 755 | | | | _ | 760 | | | | | 765 | | | |
| | | | | | | | | | | | | | | | | |
| ata | aag | aaa | aat | gag | att | ttc | cct | ttc | gat | acc | act | ata | tta | aat | gga | 2474 |
| | | | | | | | | | - | Thr | | | | | | |
| | • | 770 | | | | | 775 | | - | | | 780 | | | - | |
| | | | | | | | | | | | | | | | | |
| aat | att | cat | aaq | gag | aac | aaq | ata | gaa | gaa | gag | aaa | aat | gtg | tct | tca | 2522 |
| Asn | Ile | His | Lys | Glu | Asn | Lys | Ile | Glu | Glu | Glu | Lys | Asn | Val | Ser | Ser | |
| | 785 | | - | | | 790 | | | | | 795 | | | | | |
| | | | | | | | | | | | | | | | | |
| tct | aca | aag | tat | gat | gta | aat | aat | aag | aat | aat | aaa | aat | aat | gat | aat | 2570 |
| Ser | Thr | Lys | Tyr | Asp | Val | Asn | Asn | Lys | Asn | Asn | Lys | Asn | Asn | Asp | Asn | |
| 800 | | - | - | • | 805 | | | - | | 810 | _ | | | - | 815 | |
| | | | | | | | | | | | | | | | | |
| aqt | gaa | att | ata | aaa | tat | gaa | gat | atq | ttt | tca | aaa | qaq | acq | ttc | aca | 2618 |
| - | • | | | | | _ | - | - | | Ser | | | - | | | |
| | | | | 820 | - , | | | | 825 | | | | | 830 | | |
| | | | | | | | | | | | | | | | | |
| gat | ata | tat | aca | aat | gaa | atσ | tta | aaa | tat | tta | aad | aaa | αat | ада | aat | 2666 |
| | | | | | _ | _ | | | | Leu | _ | | - | _ | | |
| | | - 7 - | 835 | no. | 010 | 1100 | 200 | 840 | -1- | 200 | _,0 | 2,0 | 845 | 9 | | |
| | | | 033 | | | | | 040 | | | | | 045 | | | |
| ata | ata | ++- | c+ ^ | +~+ | CCC | ac+ | ato | ++= | au = | gga | tc= | uu = | tta | a++ | 222 | 2714 |
| | | | | | | • | - | | | Gly | | | - | - | | -117 |
| *** | 116 | | TEA | Sel | LTO | VIQ | | Den | OTÀ | OTÀ | ACT | 860. GTA | Leu | val | nys | |
| | | 850 | | | | | 855 | | | | | 000 | | | | |
| a = 3 | | | | | | • | | t | <u>.</u> | | <u>.</u> | | | | | 2762 |
| αLT | agt | gag | cgt | tat | cca | aat | aat | gta | cat | gat | gta | ggt | ata | gca | gaa | 2762 |

| | | | | | | | | | | 1.2 | | | | | | |
|-------|------|------|-------|-------|-------|------|------|------|-----|------|------|------|------|-----|-------|------|
| Ile | Ser | Glu | Arg | Tyr | Pro | Asn | Asn | Val | Tyr | Asp | Val | Gly | Ile | Ala | Glu | |
| | 865 | | | | | 870 | | | | | 875 | | | | | |
| | | | | | | | | | | | | | | | | |
| caa | cat | tct | gta | act | ttc | gca | gca | gct | atg | gca | atg | aat | aag | aaa | tta | 2810 |
| | | | | | Phe | | | | | | | | | | | |
| 880 | | | | | 885 | | | | | 890 | | | -10 | | 895 | |
| 000 | | | | | 003 | | | | | 030 | | | | | 093 | |
| | | | | | | | | | | | | | | | | |
| | | | | _ | ata | | | | | | | | | | - | 2858 |
| Lys | Ile | Gln | Leu | Cys | Ile | Tyr | Ser | Thr | Phe | Leu | Gln | Arg | Ala | Tyr | Asp | |
| | | | | 900 | | • | | | 905 | | | | | 910 | | |
| | | | | | | | | | | | | | | | | |
| caa | att | ata | cat | gat | ctt | aat | tta | caa | aat | ata, | cct | tta | aag | gtt | ata | 2906 |
| Gln | Ile | Ile | His | Asp | Leu | Asn | Leu | Gln | Asn | Ile | Pro | Leu | Lys | Val | Ile | |
| | | | 915 | | | | | 920 | | | | | 925 | | | |
| | | | | | | | | | | | | | | | | |
| att | gga | aga | agt | gga | tta | gta | gga | gag | gat | ggg | gca | aca | cat | caa | ggt | 2954 |
| Ile | Gly | Arg | Ser | Gly | Leu | Val | Gly | Glu | Asp | Gly | Ala | Thr | His | Gln | Gly | |
| | | 930 | | _ | | | 935 | | | | | 940 | | | | · |
| | | | | | | | | | | | | | | | | |
| ata | tat | gat | tta | tct | tat | ctt | aaa | aca | ctt | aac | aat | qca | tat | ata | ata | 3002 |
| | | | | | Tyr | | | | | | | | | | | |
| | 945 | -101 | 200 | 001 | - , - | 950 | 0-7 | | | | 955 | | -,1- | | | |
| | 743 | | | | | 950 | | | | | 223 | | | | | |
| t 0.t | ~~~ | ~~* | | | | | ++~ | 222 | 2~2 | aat | a++ | 200 | | aat | + > + | 3050 |
| | | | | | gtt | - | | | | | | | | | | 2020 |
| | PIO | ser | ASN | GIn | Val | Asp | red | туу | Arg | | Leu | Arg | rne | ATA | | |
| 960 | | | | | 965 | | | | | 970 | | | | | 975 | |
| | | | | | | | | | | | | | | | | |
| | _ | _ | - | | tct | | | | - | | | _ | | | | 3098 |
| Leu | Asp | Lys | Asp | His | Ser | Val | Tyr | Ile | Arg | Ile | Pro | Arg | Met | Asn | Ile | |
| | | | | 980 | | | | | 985 | | | | | 990 | | |
| | | | | | | | | | | | | | | | | |
| tta | agt | gat | aag | tac | atg | aaa | gga | tat | ttg | aac | att | cat | atg | aaa | aat | 3146 |
| Leu | Ser | Asp | Lys | Tyr | Met | Lys | Gly | Tyr | Leu | Asn | Ile | His | Met | Lys | Asn | |
| | | | 995 | | | | : | 1000 | | | | 1 | 1005 | | | |
| | | | | | | | | | | | | | | | | |
| gag | agc | aaa | aat | atc | gat | gta | aac | gtg | gat | ata | aac | gat | gat | gta | gat | 3194 |
| | | | | | Asp | | | | | | | | | | | |
| | | 1010 | | | • | | 1015 | | - | | | .020 | • | | - | |
| | | | | | | | | | | | | | | | | |
| 222 | tat | ant | (12.2 | W = 5 | tat | ato | gac | gat | gat | aat | +++ | ata | 222 | tea | ttt | 3242 |
| | | | | | | | | | | | | | | | | |
| | _ | Je! | GIU | GIU | Tyr | | vəħ | vəħ | voh | | | 116 | nya | SET | | |
| | 1025 | | | | | 1030 | | | |] | 1035 | | | | | |

| | | 14 | • | |
|--------------------|-----------------|---------------|---------------------|------|
| att gga aaa tct aq | ga att att aaa | atg gat aat g | aa aat aat aat aca | 3290 |
| Ile Gly Lys Ser A | g Ile Ile Lys | Met Asp Asn G | Slu Asn Asn Asn Thr | |
| 1040 | 1045 | 1050 | 1055 | |
| · | | | | |
| aat gaa cat tat to | a agc aga gga | gat aca cag a | ca aaa aaa aaa aaa | 3338 |
| Asn Glu His Tyr Se | er Ser Arg Gly | Asp Thr Gln T | hr Lys Lys Lys Lys | |
| 106 | 50 | 1065 | 1070 | |
| | | | | |
| gtt tgt atc ttt aa | c atg ggt agt | atg ctt ttt a | at gta att aat gct | 3386 |
| Val Cys Ile Phe As | n Met Gly Ser | Met Leu Phe A | sn Val Ile Asn Ala | |
| 1075 | 1 | 080 | 1085 | |
| | | | | |
| ata aaa gaa att ga | a aaa gaa caa | tat att tca c | at aat tat tot ttt | 3434 |
| | _ | | is Asn Tyr Ser Phe | |
| 1090 | 1095 | -, | 1100 | |
| | 2072 | | 1100 | |
| tca att oft oat at | a ata ttt tta | aat oot tta o | at aaa aat atg ata | 3482 |
| | | • | sp Lys Asn Met Ile | 3402 |
| 1105 | 1110 | | | |
| 1103 | 1110 | 11 | 15 | |
| | | | | 2520 |
| | | | ta att act tat gaa | 3530 |
| _ | _ | _ | eu Ile Thr Tyr Glu | |
| 1120 | 1125 | 1130 | 1135 | |
| | | | | |
| gat aat act ata go | t ggt ttt tct | aca cat ttc a | at aat tat tta ata | 3578 |
| Asp Asn Thr Ile Gl | y Gly Phe Ser | Thr His Phe A | sn Asn Tyr Leu Ile | |
| 114 | 0 | 1145 | 1150 | |
| | | | | |
| gaa aat aat tat at | t aca aaa cat | aac tta tat g | tt cat aat att tat | 3626 |
| Glu Asn Asn Tyr Il | e Thr Lys His | Asn Leu Tyr V | al His Asn Ile Tyr | |
| 1155 | 1 | 160 | 1165 | |
| | | | | |
| tta tct aat gag co | a att gaa cat | gca tct ttt a | ag gat caa caa gaa | 3674 |
| Leu Ser Asn Glu Pr | o Ile Glu His | Ala Ser Phe L | ys Asp Gln Gln Glu | |
| 1170 | 1175 | | 1180 | |
| | | • | | |
| gtc gtc aaa atg ga | t aaa tgt agt | ctt gtc aat a | ga att aaa aat tat | 3722 |
| Val Val Lys Met As | - | _ | - | |
| 1185 | 1190 | 11: | | |
| | | 4 4. | - - | |
| ctt aaa aat aat co | et aca tostotas | da taaatatata | tttctaaaat | 3770 |
| cit and dat dat CC | .c aca iyalyidd | ya cadalaldid | CCCCGGGGG | 3110 |

Leu Lys Asn Asn Pro Thr

1200 1205

tattttttt ttatacttta atgtgtacaa taaaatatat atctaaatat attttatttg 3830

tacgcttttt ttttttttt tttaattgtt atttttgtat at

3872

<210> 4

<211> 1205

<212> PRT

<213> Plasmodium falciparum

<400> 4

Met Ile Phe Asn Tyr Val Phe Phe Lys Asn Phe Val Pro Val Val Leu
1 5 10 15

Tyr Ile Leu Leu Ile Ile Tyr Ile Asn Leu Asn Gly Met Asn Asn Lys
20 25 30

Asn Gln Ile Lys Thr Glu Lys Ile Tyr Ile Lys Lys Leu Asn Arg Leu
35 40 45

Ser Arg Lys Asn Ser Leu Cys Ser Ser Lys Asn Lys Ile Ala Cys Leu 50 55 60

Phe Asp Ile Gly Asn Asp Asp Asn Arg Asn Thr Thr Tyr Gly Tyr Asn 65 70 75 80

Val Asn Val Lys Asn Asp Asp Ile Asn Ser Leu Leu Lys Asn Asn Tyr 85 90 95

Ser Asn Lys Leu Tyr Met Asp Lys Arg Lys Asn Ile Asn Asn Val Ile 100 105 110

Ser Thr Asn Lys Ile Ser Gly Ser Ile Ser Asn Ile Cys Ser Arg Asn 115 120 125

Gln Lys Glu Asn Glu Gln Lys Arg Asn Lys Gln Arg Cys Leu Thr Gln 130 135 140

Cys His Thr Tyr Asn Met Ser His Glu Gln Asp Lys Leu Ala Asn Asp

150

145

160

Asn Asn Arg Asn Asn Lys Lys Asn Phe Asn Leu Leu Phe Ile Asn Tyr 165 170 175

Phe Asn Leu Lys Arg Met Lys Asn Ser Leu Leu Asn Lys Asp Asn Phe 180 185 190

Phe Tyr Cys Lys Glu Lys Lys Leu Ser Phe Leu His Lys Ala Tyr Lys 195 200 205

Lys Lys Asn Cys Thr Phe Gln Asn Tyr Ser Leu Lys Arg Lys Ser Asn 210 215 220

Arg Asp Ser His Lys Leu Phe Ser Gly Glu Phe Asp Asp Tyr Thr Asn 225 230 235 240

Asn Asn Ala Leu Tyr Glu Ser Glu Lys Lys Glu Tyr Ile Thr Leu Asn 245 250 255

Asn Asn Asn Lys Asn Asn Asn Asn Lys Asn Asn Asp Asn Lys Asn Asn 260 265 270

Asp Asn Asp Tyr Asn Asn Asn Ser Cys Asn Asn Leu Gly Glu 275 280 285

Arg Ser Asn His Tyr Asp Asn Tyr Gly Gly Asp Asn Asn Asn Pro Cys 290 295 300

Asn Asn Asn Asn Asp Lys Tyr Asp Ile Gly Lys Tyr Phe Lys Gln Ile 305 310 315 320

Asn Thr Phe Ile Asn Ile Asp Glu Tyr Lys Thr Ile Tyr Gly Asp Glu 325 330 335

Ile Tyr Lys Glu Ile Tyr Glu Leu Tyr Val Glu Arg Asn Ile Pro Glu 340 345 350

Tyr Tyr Glu Arg Lys Tyr Phe Ser Glu Asp Ile Lys Lys Ser Val Leu 355 360 365

Phe Asp Ile Asp Lys Tyr Asn Asp Val Glu Phe Glu Lys Ala Ile Lys

Glu Glu Phe Ile Asn Asn Gly Val Tyr Ile Asn Asn Ile Asp Asn Thr Tyr Tyr Lys Lys Glu Asn Ile Leu Ile Met Lys Lys Ile Leu His Tyr Phe Pro Leu Leu Lys Leu Ile Asn Asn Pro Ser Asp Leu Lys Lys Leu Lys Lys Gln Tyr Leu Pro Leu Leu Ala His Glu Leu Lys Ile Phe Leu Phe Phe Ile Val Asn Ile Thr Gly Gly His Phe Ser Ser Val Leu Ser Ser Leu Glu Ile Gln Leu Leu Leu Leu Tyr Ile Phe Asn Gln Pro Tyr Asp Asn Val Ile Tyr Asp Ile Gly His Gln Ala Tyr Val His Lys Ile Leu Thr Gly Arg Lys Leu Leu Phe Leu Ser Leu Arg Asn Lys Lys Gly Ile Ser Gly Phe Leu Asn Ile Phe Glu Ser Ile Tyr Asp Lys Phe Gly Ala Gly His Ser Ser Thr Ser Leu Ser Ala Ile Gln Gly Tyr Tyr Glu Ala Glu Trp Gln Val Lys Asn Lys Glu Lys Tyr Gly Asn Gly Asp Ile Glu Ile Ser Asp Asn Ala Asn Val Thr Asn Asn Glu Arg Ile Phe Gln Lys Gly Ile His Asn Asp Asn Asn Ile Asn Asn Ile Asn Asn

Asn Tyr Ile Asn Pro Ser Asp Val Val Gly Arg Glu Asn Thr Asn Val

Pro Asn Val Arg Asn Asp Asn His Asn Val Asp Lys Val His Ile Ala Ile Ile Gly Asp Gly Gly Leu Thr Gly Gly Met Ala Leu Glu Ala Leu Asn Tyr Ile Ser Phe Leu Asn Ser Lys Ile Leu Ile Ile Tyr Asn Asp Asn Gly Gln Val Ser Leu Pro Thr Asn Ala Val Ser Ile Ser Gly Asn Arg Pro Ile Gly Ser Ile Ser Asp His Leu His Tyr Phe Val Ser Asn Ile Glu Ala Asn Ala Gly Asp Asn Lys Leu Ser Lys Asn Ala Lys Glu Asn Asn Ile Phe Glu Asn Leu Asn Tyr Asp Tyr Ile Gly Val Val Asn Gly Asn Asn Thr Glu Glu Leu Phe Lys Val Leu Asn Asn Ile Lys Glu Asn Lys Leu Lys Arg Ala Thr Val Leu His Val Arg Thr Lys Lys Ser Asn Asp Phe Ile Asn Ser Lys Ser Pro Ile Ser Ile Leu His Ser Ile Lys Lys Asn Glu Ile Phe Pro Phe Asp Thr Thr Ile Leu Asn Gly Asn Ile His Lys Glu Asn Lys Ile Glu Glu Glu Lys Asn Val Ser Ser Ser Thr Lys Tyr Asp Val Asn Asn Lys Asn Asn Lys Asn Asn Asp Asn Ser

Glu Ile Ile Lys Tyr Glu Asp Met Phe Ser Lys Glu Thr Phe Thr Asp

820

825

830 .

| Ile | Tyr | Thr 835 | Asn | Glu | Met | Leu | Ļys 840 | Tyr | Leu | Lys | Lys | Asp 845 | Arg | Asn | Ile |
|------------|-------------|------------|------------|------------|-------------|-------------|-------------|------------|------------|-------------|-------------|-------------|------------|------------|-------------|
| Ile | Phe 850 | Leu | Ser | Pro | Ala | Met 855 | Leu | Gly | Gly | Ser | Gly 860 | Leu | Val | Lys | Ile |
| Ser 865 | Glu | Arg | Tyr | Pro | Asn 870 | Asn | Val | Tyr | Asp | Val 875 | Gly | Ile | Ala | Glu | Gln 880 |
| His | Ser | Val | Thr | Phe 885 | Ala | Ala | Ala | Met | Ala 890 | Met | Asn | Lys | Lys | Leu 895 | Lys |
| Ile | Gln | Leu | Cys 900 | Ile | Tyr | Ser | Thr | Phe 905 | Leu | Gln | Arg | Ala | Tyr 910 | Asp | Gln |
| Ile | Ile | His 915 | Asp | Leu | Asn | Leu | Gln 920 | Asn | Ile | Pro | Leu | Lys 925 | Val | Ile | Ile |
| Gly | Arg 930 | Ser | Gly | Leu | Val | Gly 935 | Glu | Asp | Gly | Ala | Thr 940 | His | Gln | Gly | Ile |
| Tyr 945 | Asp | Leu | Ser | Tyr | Leu 950 | Gly | Thr | Leu | Asn | Asn 955 | Ala | Tyr | Ile | Ile | Ser 960 |
| Pro | Ser | Asn | Gln | Val 965 | Asp | Leu | Lys | Arg | Ala 970 | Leu | Arg | Phe | Ala | Tyr 975 | Leu |
| Asp | Lys | Asp | His 980 | Ser | Val | Tyr | Ile | Arg 985 | Ile | Pro | Arg | Met | Asn 990 | Ile | Leu |
| Ser | Asp | Lys 995 | Tyr | Met | Lys | - | Туг 1000 | Leu | Asn | Ile | | Met .005 | Lys | Asn | Glu |
| | Lys 1010 | Asn | Ile | Asp | Val | Asn 1015 | Val | Asp | Ile | | Asp 1020 | Asp | Val | Asp | Lys |
| Tyr 025 | Ser | Glu | Glu | _ | Met 1030 | Asp | Asp | Asp | | Phe 1035 | Ile | Lys | Ser | | Ile .040 |
| Gly | Lys | Ser | Arg | Ile | Ile | Lys | Met | Asp | Asn | Glu | Asn | Asn | Asn | Thr | Asn |

1045

1050

1055

Glu His Tyr Ser Ser Arg Gly Asp Thr Gln Thr Lys Lys Lys Val $1060 \hspace{1.5cm} 1065 \hspace{1.5cm} 1070$

Cys Ile Phe Asn Met Gly Ser Met Leu Phe Asn Val Ile Asn Ala Ile 1075 1080 1085

Lys Glu Ile Glu Lys Glu Gln Tyr Ile Ser His Asn Tyr Ser Phe Ser 1090 1095 1100

Ile Val Asp Met Ile Phe Leu Asn Pro Leu Asp Lys Asn Met Ile Asp 105 1110 1115 1120

His Val Ile Lys Gln Asn Lys His Gln Tyr Leu Ile Thr Tyr Glu Asp 1125 1130 1135

Asn Thr Ile Gly Gly Phe Ser Thr His Phe Asn Asn Tyr Leu Ile Glu 1140 1145 1150

Asn Asn Tyr Ile Thr Lys His Asn Leu Tyr Val His Asn Ile Tyr Leu 1155 1160 1165

Ser Asn Glu Pro Ile Glu His Ala Ser Phe Lys Asp Gln Gln Glu Val 1170 1175 1180

Val Lys Met Asp Lys Cys Ser Leu Val Asn Arg Ile Lys Asn Tyr Leu 185 1190 1195 1200

Lys Asn Asn Pro Thr 1205

<210> 5

<211> 3147

<212> DNA

<213> Plasmodium falciparum

<220>

<221> CDS

<222> (199)..(2670)

| | | | | | | | | | | ~ L | | | | | | |
|------|-------|----------|-------|-------|-------|------|-------|-------|-------|------------|-----|------|-----|------|--------|-----|
| <400 |)> 5 | | | | | | | | | | | | | | | |
| ttt | catt | tt (| cttta | CCC | ac at | tata | tata | t ata | atata | atat | aat | atat | ata | tata | atatta | 60 |
| tata | atttç | gat a | atato | gatti | ta aa | aatt | gtaad | c ata | aaaa | aaaa | taa | ttat | att | aaat | atgtgt | 120 |
| atad | catct | cc a | aacat | ataa | aa ta | atta | tttt | t tai | ttati | tatt | ttt | tttt | ttt | tttt | tcataa | 180 |
| tgc | ctgaa | ata a | accad | caaa | atg | agt | tat | ata | aaa | aga | ctg | att | ctt | ttt | atg | 231 |
| | | | | | Met | Ser | Tyr | Ile | Lys | Arg | Leu | Ile | Leu | Phe | Met | |
| | | | | | 1 | | | | 5 | | | | | 10 | | |
| tta | cta | * | tat | tct | cat | gta. | aaa | att | 222 | aaa | tta | ttt | att | aaa | att | 279 |
| | - | | | | | _ | | | | | | | | Lys | | |
| u | 200 | 1110 | 15 | 501 | | ,,,, | -,- | 20 | -,- | -,- | | | 25 | , . | -10 | |
| | | | | | | | | | | | | | | | | |
| tst | aat | gta | aac | ata | ttt | ttt | gca | gaa | gca | aag | aaa | aat | gga | aaa | aag | 327 |
| Ser | Asn | Val | Asn | Ile | Phe | Phe | Ala | Glu | Ala | Lys | Lys | Asn | Gly | Lys | Lys | |
| | | 30 | | | | | 35 | | | | | 40 | | | | |
| | | | | | | | | | | | | | | | | |
| gaa | ttc | ttt | ctt | ttt | tta | cta | aat | ata | aaa | aaa | aat | agc | caa | cag | aaa | 375 |
| Glu | Phe | Phe | Leu | Phe | Leu | Leu | Asn | Ile | Lys | Lys | Asn | Ser | Gln | Gln | Lys | |
| | 45 | | | | | 50 | | | | | 55 | | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | gat | | 423 |
| | Thr | Tyr | His | Ile | | Lys | Arg | Asn | Thr | | Asn | Lys | Ser | Asp | | |
| 60 | | | | | 65 | | | | | 70 | | | | | 75 | |
| tta | tat | tct | tta | cta | aat | gaa | gaa | aaa | aat | tct | tca | aaa | aaq | gaa | tat | 471 |
| | | | | | | | | | | | | | | Glu | | |
| | _ | | | 80 | | | | _ | 85 | | | - | _ | 90 | | |
| | | | | * | | | | | | | | | | | | |
| aaa | aat | tta | aaa | gat | gaa | gaa | aaa | tat | aat | atc | ata | caa | aat | ata | aaa | 519 |
| Lys | Asn | Leu | Lys | Asp | Glu | Glu | Lys | Tyr | Asn | Ile | Ile | Gln | Asn | Ile | Lys | |
| | | | 95 | | | | | 100 | | | | | 105 | | | |
| | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | gaa · | 567 |
| Lys | Tyr | Cys | Glu | Cys | Thr | Lys | Lys | Tyr | Lys | Arg | Leu | Pro | Thr | Arg | Glu | |
| | | 110 | | | | | 115 | | | | | 120 | | | | |

Val Val Ile Gly Asn Val Lys Ile Gly Gly Asn Asn Lys Ile Ala Ile

PCT/EP99/07055

caa act atg gct agc tgt gat aca aga aat gta gaa gaa tgt gta tat Gln Thr Met Ala Ser Cys Asp Thr Arg Asn Val Glu Glu Cys Val Tyr caa att aga aaa tgt aaa gat ttg ggt gct gac att gta agg ttg act Gln Ile Arg Lys Cys Lys Asp Leu Gly Ala Asp Ile Val Arg Leu Thr gtt caa gga gtt caa gaa gca caa gct agt tat cat att aaa gaa aaa Val Gln Gly Val Gln Glu Ala Gln Ala Ser Tyr His Ile Lys Glu Lys tta tta tct gaa aat gta aat atc cca tta gta gca gat att cat ttt Leu Leu Ser Glu Asn Val Asn Ile Pro Leu Val Ala Asp Ile His Phe aat cct aaa ata gct tta atg gca gct gat gtg ttt gaa aaa att cga Asn Pro Lys Ile Ala Leu Met Ala Ala Asp Val Phe Glu Lys Ile Arg gtg aat cca gga aat tat gtt gat gga aga aaa aaa tgg ata gat aaa Val Asn Pro Gly Asn Tyr Val Asp Gly Arg Lys Lys Trp Ile Asp Lys Val Tyr Lys Thr Lys Glu Glu Phe Asp Glu Gly Lys Leu Phe Ile Lys gaa aaa ttt gta cca tta att gaa aaa tgt aaa aga tta aat aga gca Glu Lys Phe Val Pro Leu Ile Glu Lys Cys Lys Arg Leu Asn Arg Ala ata aga att gga aca aat cat gga tcc ctt tca tct cga gta tta tca Ile Arg Ile Gly Thr Asn His Gly Ser Leu Ser Ser Arg Val Leu Ser tat tat gga gat aca cca tta ggt atg gta gaa tcg gct ttt gag ttt Tyr Tyr Gly Asp Thr Pro Leu Gly Met Val Glu Ser Ala Phe Glu Phe

| | | | | | | | | | 2 | 23 | | | | | | |
|-----|------|-----|-------|-----|-----|-----|------|-----|-----|-----|------|-------------|-------|-----|-------------|------|
| tct | gat | tta | tgt | att | gaa | aac | aat | ttt | tac | aat | ctt | gtt | ttt | tct | atg | 1143 |
| Ser | Asp | Leu | Cys | Ile | Glu | Asn | Asn | Phe | Tyr | Asn | Leu | Val | Phe | Ser | Met | |
| 300 | | | | | 305 | | | | | 310 | | _ | | | 315 | |
| | | | | | | | | | | | • | | | | | |
| aaa | act | tct | aat | act | tat | att | atα | ata | caa | tct | tat | aga | tta | tta | ota | 1191 |
| | | | | - | | - | | | | | | | | | | |
| гуs | AIA | Ser | ASII | | ıyı | Val | Met | 116 | | Ser | Tyr | ALG | теп | | Val | |
| | | | | 320 | | | | | 325 | | | | | 330 | | |
| | | | | | | | | | | | | | | | | |
| tct | aaa | caa | tat | gaa | aga | aat | atg | atg | ttc | cct | ata | cat | tta | gga | gtt | 1239 |
| Ser | Lys | Gln | Tyr | Glu | Arg | Asn | Met | Met | Phe | Pro | Ile | His | Leu | Gly | Val | • |
| | | | 335 | | | | | 340 | | | | | 345 | | | |
| | | | | | | | | | | | | | | | | |
| aca | gaa | qca | gga | ttt | aat | gat | aat | gga | aga | ata | aaa | tct | tat | tta | qqt | 1287 |
| | - | - | | | | - | | | | | Lys | | | | | |
| | | 350 | CLJ | | O.J | | 355 | , | 9 | | -,- | 360 | - 1,- | | ~_ 1 | |
| | | 330 | | | | | 333 | | | | | 300 | | | | |
| | | | | | | | | | | | | | | | | 1005 |
| | | | | | | - | | | | | acc | | - | | | 1335 |
| Ile | Gly | Ser | Leu | Leu | Tyr | Asp | Gly | Ile | Gly | Asp | Thr | Ile | Arg | Ile | Ser | |
| | 365 | | | | | 370 | | | | | 375 | | | | | |
| | | | | | | | | | | | | | | | | |
| tta | aca | gaa | gat | cct | tgg | gaa | gag | tta | act | cct | tgt | aaa | aaa | tta | gtt | 1383 |
| Leu | Thr | Glu | Asp | Pro | Trp | Glu | Glu | Leu | Thr | Pro | Cys | Lys | Lys | Leu | Val | |
| 380 | | | | | 385 | | | | | 390 | | | | | 395 | |
| | | | | | | | | | | | | | | | | |
| gaa | aat | tta | 220 | 222 | ana | ata | ttt | tat | aat | ааа | aat | t ++ | 222 | gaa | gat | 1431 |
| | | | - | | - | | | | | | Asn | | | | | |
| GIU | POII | rea | rås | - | Arg | 116 | FIIG | ıyı | | GIU | N311 | riie | Буз | | nap | |
| | | | | 400 | | | | | 405 | | | | | 410 | | |
| | | | | | | | | | | | | | | | | |
| aat | gaa | tta | aaa | aat | aat | gaa | atg | gat | acc | aaa | aat | cta | tta | aat | ttt | 1479 |
| Asn | Glu | Leu | Lys | Asn | Asn | Glu | Met | Asp | Thr | Lys | Asn | Leu | Leu | Asn | Phe | |
| | | | 415 | | | | | 420 | | | | _ | 425 | | | |
| | | | | | | | | | | | | | | | | |
| gaa | gaa | aat | tat | cga | aat | ttt | aat | aat | ata | aaa | aaa | aga | aat | gta | gaa | 1527 |
| | | | | _ | | | | | | | Lys | | | | | |
| | | 430 | - , - | 3 | | | 435 | | | - | • | 440 | | | | |
| | | 430 | | | | | 455 | | | | | 440 | | | | |
| | | | | | | | | | | | | | | | | |
| | | | | - | | | | | | | | | | | gta · | 1575 |
| Lys | Asn | Asn | Asn | Val | Leu | His | Glu | Glu | Cys | Thr | Ile | Gly | Asn | Val | Val | |
| | 445 | | | | | 450 | | | | | 455 | | | | | |
| | | | | | | | | | | | | | | | | |
| acc | ata | aaa | gag | tta | gaa | gat | tct | ctg | caa | att | ttt | aaa | gat | tta | aat | 1623 |
| | | | | | | | | | | | Phe | | | | | |
| | | 4 - | | | | • | | | | | | _ | - | | | |

tta gaa gta gat tca aat gga aat ttg aaa aag gga gcc aaa aca act Leu Glu Val Asp Ser Asn Gly Asn Leu Lys Lys Gly Ala Lys Thr Thr gat atg gtt att ata aat gat ttt cat aat ata aca aat tta gga aaa Asp Met Val Ile Ile Asn Asp Phe His Asn Ile Thr Asn Leu Gly Lys aaa act gtg gat aaa tta atg caa gtg gga att aat ata gta gtt caa Lys Thr Val Asp Lys Leu Met Gln Val Gly Ile Asn Ile Val Val Gln tat gaa cca cat aat ata gaa ttt ata gaa aaa atg gaa cca aat aat Tyr Glu Pro His Asn Ile Glu Phe Ile Glu Lys Met Glu Pro Asn Asn gat aat aat aat aat aat aat aat aat ata tta ttt tat gtg gat Asp Asn Asn Asn Asn Asn Asn Asn Asn Ile Leu Phe Tyr Val Asp ata aaa aat att atg aac agt tca gaa aaa aat att aaa tta agt aat Ile Lys Asn Ile Met Asn Ser Ser Glu Lys Asn Ile Lys Leu Ser Asn tct aaa gga tat gga tta att tta aac gga aaa gaa gat ata caa acc Ser Lys Gly Tyr Gly Leu Ile Leu Asn Gly Lys Glu Asp Ile Gln Thr ata aaa aaa ata aaa gaa tta aat cgt cgt cct tta ttc att cta tta Ile Lys Lys Ile Lys Glu Leu Asn Arg Arg Pro Leu Phe Ile Leu Leu aaa tca gat aac ata tat gaa cat gta tta ata acc aga aga att aat Lys Ser Asp Asn Ile Tyr Glu His Val Leu Ile Thr Arg Arg Ile Asn gaa ctt tta caa tcc tta aat ata aat ata cct tat ata cat tat gtt Glu Leu Leu Gln Ser Leu Asn Ile Asn Ile Pro Tyr Ile His Tyr Val

| • | wo | 00/17 | 7233 | | | | | | | | | | | | | PCT/EP99/07055 |
|--------|-----|-------|------|---------|----------|-----|-----|------|------|-----------|-----|-----|------|-----|-------------|----------------|
| gat | att | aat | tca | aac | aat | tat | gat | gat | ata | 25 tta | att | aat | t.ca | aca | tta | 2151 |
| - | | | | | Asn | | • | • | | | • | | | | _ | |
| • | | | | 640 | | • | • | • | 645 | • | | | | 650 | | - |
| | | | | • | | | | | | | | • | • | | | |
| tat | gca | gga | agt | tgt | ttg | atg | gat | tta | atg | ggg | gat | ggt | ctt | att | gtt | 2199 |
| Tyr | Ala | Gly | Ser | Cys | Leu | Met | Asp | Leu | Met | Gly | Asp | Gly | Leu | Ile | Val | |
| | | | 655 | | | | | 660 | | | | | 665 | | | |
| | | | | | | | | | | | | | | | | |
| aac | gta | act | aat | gat | gtt | ctt | aca | aat | aaa | aaa | aag | ata | gaa | aca | aaa | 2247 |
| Asn | Val | Thr | Asn | Asp | Val | Leu | Thr | Asn. | Lys | Lys | Lys | Ile | Glu | Thr | Lys | |
| | | 670 | | | | | 675 | | | | | 680 | | | | |
| | | | | | | | | | | | | | | | | |
| tat | gat | gaa | aaa | gaa | gaa | gta | gag | gaa | gag | gga | aac | aat | aaa | gat | att | 2295 |
| Tyr | Asp | Glu | Lys | Glu | Glu | Val | Glu | Glu | Glu | Gly | Asn | Asn | Lys | Asp | Ile | |
| | 685 | | | | | 690 | | | | | 695 | | | | | |
| | | | | | | | | | | | | | | | | |
| cat | aga | ctt | ttg | agc | aga | gtt | gca | tta | aat | tca | ttt | tta | aca | tta | aat | 2343 |
| His | Arg | Leu | Leu | Ser | Arg | Val | Ala | Leu | Asn | Ser | Phe | Leu | Thr | Leu | Asn | |
| 700 | | | | | 705 | | | | | 710 | | | | | 715 | |
| • | | | | | | | | | | | | | | | | |
| | | | • | | aga | | - | | | | | • | | | • | 2391 |
| Ile | Leu | Gln | Asp | | Arg | Ile | Arg | Leu | | Lys | Thr | Asp | Tyr | | Ala | |
| | | | | 720 | | | | | 725 | | | | | 730 | | |
| | | | | | | | | | | | | | | | | 0.420 |
| , | | | | | aga - | | | | | | | | | | | 2439 |
| Cys | Pro | ser | - | GIĀ | Arg | Thr | Leu | | Asn | 116 | GIN | Glu | | Thr | Lys | |
| | | | 735 | | | | | 740 | | | ٠ | | 745 | | | |
| 223 | | at~ | | | | | 050 | ++- | 222 | ~~~ | ~++ | 2 | 2+4 | ~~~ | ~+ ~ | 2407 |
| | | _ | | | aca | | | | | | · . | | | _ | _ | 2487 |
| כענ | 116 | 750 | ոչ | nea | Thr | GIÀ | | nea | пÃа | arl | AaT | _ | 116 | ura | val | |
| | | 730 | | | | | 755 | | | | | 760 | | | | |
| ato | ana | tot | 2++ | at+ | aat | aa+ | ata | aas | ma a | atm | ac= | na+ | ac. | cat | +++ | 2535 |
| | | | | • | Asn | | | | | | - | - | - | | | 2333 |
| 1.1¢ C | grā | cys | TTG | val | ASN | GTÀ | тте | GTÅ | GIU | MEC | wrg | vab | WTG | urz | rue | |

gag tta gta gaa aga aat ata cct gag gaa gaa gct tgt gat aaa ttg 2631 Glu Leu Val Glu Arg Asn Ile Pro Glu Glu Glu Ala Cys Asp Lys Leu

ggt tat gtt ggt agt gca cct aaa aaa att gat tta tat tat ggt aaa Gly Tyr Val Gly Ser Ala Pro Lys Lys Ile Asp Leu Tyr Tyr Gly Lys

800

805 26

ata gaa tta att aaa aaa cat aac aaa tgg aaa gat cca taaattgaat 2680 Ile Glu Leu Ile Lys Lys His Asn Lys Trp Lys Asp Pro 815 820

<210> 6

<211> 824

<212> PRT -

<213> Plasmodium falciparum

<400> 6

Met Ser Tyr Ile Lys Arg Leu Ile Leu Phe Met Leu Leu Phe Tyr Ser

1 5 10 15

His Val Lys Ile Lys Lys Leu Phe Ile Lys Ile Ser Asn Val Asn Ile
20 25 30

Phe Phe Ala Glu Ala Lys Lys Asn Gly Lys Lys Glu Phe Phe Leu Phe
35 40 45

Leu Leu Asn Ile Lys Lys Asn Ser Gln Gln Lys Lys Thr Tyr His Ile
50 55 60

Thr Lys Arg Asn Thr Ile Asn Lys Ser Asp Phe Leu Tyr Ser Leu Leu

Asn Glu Glu Gly Asn Ser Ser Lys Lys Glu Tyr Lys Asn Leu Lys Asp Glu Glu Lys Tyr Asn Ile Ile Gln Asn Ile Lys Lys Tyr Cys Glu Cys Thr Lys Lys Tyr Lys Arg Leu Pro Thr Arg Glu Val Val Ile Gly Asn Val Lys Ile Gly Gly Asn Asn Lys Ile Ala Ile Gln Thr Met Ala Ser Cys Asp Thr Arg Asn Val Glu Glu Cys Val Tyr Gln Ile Arg Lys Cys Lys Asp Leu Gly Ala Asp Ile Val Arg Leu Thr Val Gln Gly Val Gln Glu Ala Gln Ala Ser Tyr His Ile Lys Glu Lys Leu Leu Ser Glu Asn Val Asn Ile Pro Leu Val Ala Asp Ile His Phe Asn Pro Lys Ile Ala Leu Met Ala Ala Asp Val Phe Glu Lys Ile Arg Val Asn Pro Gly Asn Tyr Val Asp Gly Arg Lys Lys Trp Ile Asp Lys Val Tyr Lys Thr Lys Glu Glu Phe Asp Glu Gly Lys Leu Phe Ile Lys Glu Lys Phe Val Pro Leu Ile Glu Lys Cys Lys Arg Leu Asn Arg Ala Ile Arg Ile Gly Thr

Pro Leu Gly Met Val Glu Ser Ala Phe Glu Phe Ser Asp Leu Cys Ile

Asn His Gly Ser Leu Ser Ser Arg Val Leu Ser Tyr Tyr Gly Asp Thr

Glu Asn Asn Phe Tyr Asn Leu Val Phe Ser Met Lys Ala Ser Asn Ala Tyr Val Met Ile Gln Ser Tyr Arg Leu Leu Val Ser Lys Gln Tyr Glu Arg Asn Met Met Phe Pro Ile His Leu Gly Val Thr Glu Ala Gly Phe Gly Asp Asn Gly Arg Ile Lys Ser Tyr Leu Gly Ile Gly Ser Leu Leu Tyr Asp Gly Ile Gly Asp Thr Ile Arg Ile Ser Leu Thr Glu Asp Pro Trp Glu Glu Leu Thr Pro Cys Lys Leu Val Glu Asn Leu Lys Lys Arg Ile Phe Tyr Asn Glu Asn Phe Lys Glu Asp Asn Glu Leu Lys Asn Asn Glu Met Asp Thr Lys Asn Leu Leu Asn Phe Glu Glu Asn Tyr Arg Asn Phe Asn Asn Ile Lys Lys Arg Asn Val Glu Lys Asn Asn Asn Val Leu His Glu Glu Cys Thr Ile Gly Asn Val Val Thr Ile Lys Glu Leu Glu Asp Ser Leu Gln Ile Phe Lys Asp Leu Asn Leu Glu Val Asp Ser Asn Gly Asn Leu Lys Lys Gly Ala Lys Thr Thr Asp Met Val Ile Ile Asn Asp Phe His Asn Ile Thr Asn Leu Gly Lys Lys Thr Val Asp Lys

Leu Met Gln Val Gly Ile Asn Ile Val Val Gln Tyr Glu Pro His Asn

| 515 | 520 | 525 |
|-----|-----|-----|
| | | |

| | | J | | | | | 3.0 | | | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 11 | Glu 530 | Phe | Ile | Glu | Lys | Met 535 | Glu | Pro | Asn | Asn | Asp 540 | Asn | Asn | Asn | Asn |
| Asn 545 | Asn | Asn | Asn | Asn | Ile 550 | Leu | Phe | Tyr | Val | Asp 555 | Ile | Lys | Asn | Ile | Met 560 |
| Asn | Ser | Ser | Glu | Lys 565 | Asn | Ile | Lys | Leu | Ser 570 | Asn | Ser | Lys | Gly | Tyr 575 | Gly |
| Leu | Ile | Leu | Asn 580 | Gly | Lys | Glu | Asp | Ile 585 | Gln | Thr | Ile | Lys | Lys 590 | Ile | Lys |
| Glu | Leu | Asn 595 | Arg | Arg | Pro | Leu | Phe 600 | Ile | Leu | Leu | Lys | Ser 605 | Asp | Asn | Ile |
| Tyr | Glu 610 | His | Val | Leu | Ile | Thr 615 | Arg | Arg | Ile | Asn | Glu 620 | Leu | Leu | Gln | Ser |
| Leu 625 | Asn | Ile | Asn | Ile | Pro 630 | Tyr | Ile | His | Tyr | Val 635 | Asp | Ile | Asn | Ser | Asn 640 |
| Asn | Tyr | Asp | Asp | Ile 645 | Leu | Val | Asn | Ser | Thr 650 | Leu | Tyr | Ala | Gly | Ser 655 | Cys |
| Leu | Met | Asp | Leu 660 | Met | Gly | Asp | Gly | Leu 665 | Ile | Val | Asn | Val | Thr 670 | Asn | Asp |
| Val | Leu | Thr 675 | Asn | Lys | Lys | Lys | Ile 680 | G1u | Thr | Lys | Tyr | Asp 685 | Glu | Lys | Glu |
| Glu | Val 690 | Glu | Glu | Glu | Gly | Asn 695 | Asn | Lys | Asp | Ile | His 700 | Arg | Leu | Leu | Ser |
| Arg 705 | Val | Ala | Leu | Asn | Ser 710 | Phe | Leu | Thr | Leu | Asn 715 | Ile | Leu | Gln | Asp | Thr 720 |
| Arg | Ile | Arg | Leu | Phe 725 | Lÿs | Thr | Asp | Tyr | Ile 730 | Ala | Cys | Pro | Ser | Cys 735 | Gly |

Arg Thr Leu Phe Asn Ile Gln Glu Thr Thr Lys Lys Ile Met Lys Leu

WELTORGANISATION FÜR GEISTIGES EIGENTUM Internationales Büro

INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation 7:

C12N 9/90, 9/10, 9/12, C12Q 1/48

(11) Internationale Veröffentlichungsnummer: WO 00/17233

A3 (43) Internationales

Veröffentlichungsdatum:

30. März 2000 (30.03.00)

(21) Internationales Aktenzeichen:

PCT/EP99/07055

(22) Internationales Anmeldedatum:

22. September 1999

(22.09.99)

(30) Prioritätsdaten:

198 43 279.8 199 23 567.8 22. September 1998 (22.09.98) DE

21. Mai 1999 (21.05.99)

(71)(72) Anmelder und Erfinder: JOMAA, Hassan [DE/DE]; Breslauer Strasse 24, D-35398 Gießen (DE).

(74) Anwälte: PANTEN, Kirsten usw.; Reichel und Reichel, Parkstrasse 13, D-60322 Frankfurt am Main (DE).

(81) Bestimmungsstaaten: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht

Mit internationalem Recherchenbericht.

(88) Veröffentlichungsdatum des internationalen Recherchenbe-25. Mai 2000 (25.05.00) richts:

- (54) Title: GENES OF THE 1-DESOXY-D-XYLULOSE BIOSYNTHETIC PATHWAY
- (54) Bezeichnung: GENE DES 1-DESOXY-D-XYLULOSE-BIOSYNTHESEWEGS

(57) Abstract

The invention relates to the 1-desoxy- D-xylulose- 5-phosphate reductoisomerase gene, the 1-desoxy- D-xylulose- 5-phosphatesynthase gene and the gcpE gene of the 1-desoxy- D-xylulose biosynthetic pathway and to their use for transforming vectors, host organisms and plants and for determining substances that inhibit this biosynthetic pathway.

(57) Zusammenfassung

Die vorliegende Erfindung betrifft das 1-Desoxy- D-xylulose- 5-phosphatreduktoisomerase -Gen, das 1-Desoxy- D-xylulose-5-phosphat- Synthase- Gen und das gcpE-Gen des 1-Desoxy- D-xylulose- Biosynthesewegs und ihre Verwendung zur Transformation von Vektoren, Wirtsorganismen und Pflanzen und zur Bestimmung von Stoffen, die diesen Biosyntheseweg inhibieren.

LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

| | 4 May 2 | | | | _ | | |
|----|------------------------------|----|-----------------------------|-----|-----------------------------|----|------------------------|
| AL | Albanien | ES | Spanien | LS | Lesotho | SI | Slowenien |
| AM | Armenien | Fl | Finnland | LT | Litauen | SK | Slowakei |
| AT | Osterreich | FR | Frankreich | LU | Luxemburg | SN | Senegal |
| AU | Australien | GA | Gabun | LV | Lettland | SZ | Swasiland |
| AZ | Aserbaidschan | GB | Vereinigtes Königreich | MC | Monaco | TD | Tschad |
| BA | Bosnien-Herzegowina | GE | Georgien | MD | Republik Moldau | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagaskar | TJ | Tadschikistan |
| BE | Belgien | GN | Guinea | MK | Die ehemalige jugoslawische | TM | Turkmenistan |
| BF | Burkina Faso | GR | Griechenland | | Republik Mazedonien | TR | Türkei |
| BG | Bulgarien | HU | Ungam | ML | Mali | TT | Trinidad und Tobago |
| BJ | Benin | IE | Irland | MN | Mongolei | UA | Ukraine |
| BR | Brasilien | IL | Israel | MR | Mauretanien | UG | Uganda |
| BY | Belarus | IS | Esland | MW | Malawi | US | Vereinigte Staaten von |
| CA | Kanada | IT | Italien | MX | Mexiko | | Amerika |
| CF | Zentralafrikanische Republik | JP | Japan | NE | Niger | UZ | Usbekistan |
| CG | Kongo | KE | Kenia | NL | Niederlande | VN | Vietnam |
| CH | Schweiz | KG | Kirgisistan | NO | Norwegen | YU | Jugoslawien |
| CI | Côte d'Ivoire | KP | Demokratische Volksrepublik | NZ. | Neusceland | zw | Zimbabwe |
| CM | Kamerun | | Korea | PL | Polen | | |
| CN | China | KR | Republik Korea | PT | Portugal | | |
| CU | Kuba | KZ | Kasachstan | RO | Rumanien | | |
| CZ | Tschechische Republik | LC | St. Lucia | RU | Russische Föderation | | |
| DE | Deutschland | LI | Liechtenstein | SD | Sudan | | |
| DK | Dänemark | LK | Sri Lanka | SE | Schweden | | |
| EE | Estiand | LR | Liberia | SG | Singapur | | |
| | | | | | - - | | |

Int. Jonel Application No PCT/EP 99/07055

| A CLASSI | FICATION OF RUBLECT MATTER C12N9/90 C12N9/10 C12N9/12 | C1201/40 | |
|-------------|---|--|---|
| 110/ | C12n3/30 C12n3/10 C12n3/12 | . C1241/40 | |
| | | | • |
| | international Patent Classification (IPC) or to both national classific SEARCHED | ation and IPC | |
| | currentation searched (classification system followed by classificati | on symbols) | |
| IPC 7 | C12N C12Q | • | |
| | | | |
| Documentat | son searched other than minimum documentation to the extent that a | such documents are included in the fields or | beroned |
| Bectronic d | ata base consulted during the international search (name of data be | se and, where practical, search terms used |) |
| | • | | |
| | • | | |
| | | | |
| C. DOCUM | ENTS CONSIDERED TO BE RELEVANT | | |
| Category * | Challon of document, with indication, where appropriate, of the rel | evant passages | Relevant to claim No. |
| | | | |
| E,L | WO 99 52938 A (HASSAN JOMAA) 21 October 1999 (1999-10-21) | | 1,2,4, 8-12, |
| | 21 000000 1555 (1555 10 21) | | 16-18 |
| | see SeqID's; Priority of inv. sl | | , |
| | W09952938 and PCTEP99/07055 and PE19816196.4,DE19825585.3,DE1982 | | |
| | 9831637.2,DE19831639.9 or DE1983 | 1638.0 may | |
| ļ | be invalid; A.4C(4) PC. | | |
| l | | -/ | |
| 1 | | | |
| ł | | | |
| 1 | | | |
| | | | |
| İ | | | · |
| Í | | | |
| | | | |
| [V] E- | finer documents are listed in the continuation of box C. | Y Patent family members are listed | i |
| 10 | | A Date is the state of the stat | |
| 1 | stegories of cited documents : | "T" later document published after the inte- or priority date and not in conflict with | |
| const | ent defining the general state of the art which is not dered to be of particular relevance | cited to understand the principle or the invention | eary underlying the |
| filing | | "X" document of particular relevance; the cannot be considered novel or cannot | t be considered to |
| which | ent which may throw doubts on priority claim(s) or is ched to establish the publication date of another in or other special reason (as specified) | Involve an inventive step when the do "Y" document of particular relevance; the | curnent is taken eigne cialmed invention |
| O" docum | nt or other special reason (as specially) ent referring to an oral disclosure, use, exhibition or means | cannot be considered to involve an in document is combined with one or m ments, such combination being obvio | ore other such docu- |
| "P" docum | means ent published prior to the international filing date but than the priority date claimed | in the art. | _ |
| <u> </u> | actual completion of the International search | "A" document member of the same patent Date of mailing of the international se | |
| 7 | March 2000 | 23/03/2000 | |
| Name and | mailing address of the ISA | Authorized officer | |
| 1 | European Patent Office, P.B. 5818 Patentiaan 2 NL ~ 2280 HV Rijewijk | | |
| | Tel. (+31~70) 340—2040, Tx. 31 651 epo nl, Fax: (+31~70) 340—3016 | Hoekstra, S | |

ini. Jonel Application No PCT/EP 99/07055

| C/Continu | METION) DOCUMENTS CONSIDERED TO BE RELEVANT | PCT/EP 99/07055 |
|------------|---|-----------------------|
| Category * | | Relevant to claim No. |
| X | PUTRA ET AL: "Incorporation of '2,3-'13!C2!- and '2,4-'13!C2!-D-1-Deoxyxylulose into ubiquinone of Escherichia coli via the Mevalonate-Independent pathway for Isoprenoid Biosynthesis" TETRAHEDRON LETTERS,NL,ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, vol. 39, no. 39, 1998, pages 23-26-26, XP002116676 ISSN: 0040-4039 figure 1 | 15 |
| X | KUZUYAMA ET AL: "Direct formation of 2-C-Methyl-D-Erythritol 4-phosphate from 1-Deoxy-D-Xylulose 5-phosphate Reductoisomerase, a new enzyme in the non-mevalonate pathway to isopentenyl diphosphate" TETRAHEDRON LETTERS, NL, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, vol. 39, no. 39, 1998, pages 4509-4512-44512, XP002116675 ISSN: 0040-4039 figure 1 | 15 |
| Ρ,Χ | SCHWENDER, J. ET AL.: "Cloning and heterologous expression of a cDNA encoding 1-deoxy-d-xylulose-5-phosphate reductoisomerase of Arabidopsis thaliana" FEBS LETTERS, vol. 455, July 1999 (1999-07), pages 140-144, XP002132424 the whole document | 1,9-12 |
| P,A | DE 197 52 700 A (HOECHST SCHERING AGREVO GMBH) 2 June 1999 (1999-06-02) the whole document | 1–12 |
| A | LANGE ET AL: "A family of transketolases that directs isoprenoid biosynthesis via a mevalonate-independent pathway" FASEB JOURNAL, US, FED. OF AMERICAN SOC. FOR EXPERIMENTAL BIOLOGY, BETHESDA, MD, vol. 95, March 1998 (1998-03), pages 2100-2104, XP002116672 ISSN: 0892-6638 the whole document | 1-18 |
| P,X | EMINV DATABASE: "AC: EF111813" PLASMODIUM FALCIPARUM 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE, 11 January 1999 (1999-01-11), XP002132425 see : Scores | 1,9-12 |
| | -/ | |

Inh. Jonel Application No PCT/EP 99/07055

| | | PCI/EP 99 | 7 0 7 0 3 3 |
|------------|---|-----------|----------------------|
| C.(Continu | ntion) DOCUMENTS CONSIDERED TO BE RELEVANT | | |
| Category * | Citation of document, with indication, where appropriate, of the relevant passages | | Relevant to dalm No. |
| P,X | TREMBL DATABASE: "AC: 096693" PLASMODIUM FALCIPARUM 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE, 1 May 1999 (1999-05-01), XP002132426 see : Scores | | 1,9-12 |
| X | SPRENGER ET AL: "Identification of a thiamin-dependent synthase in Escherichia coli required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin, and pyridoxol" FASEB JOURNAL, US, FED. OF AMERICAN SOC. FOR EXPERIMENTAL BIOLOGY, BETHESDA, MD, vol. 94, November 1997 (1997-11), pages 12857-12862, XP002116674 ISSN: 0892-6638 figure 2 | | 2,9-12 |
| A | figure 2 figure 1 | • | 15 |
| P,X | TREMBL DATABASE: "AC: QZ8HO" CHLAMYDIA PNEUMONIAE GCPE PROTEIN, 1 May 1999 (1999-05-01), XP002132427 see : scores | | 3,9-12 |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

Information on petent family members

Inh. Sonal Application No PCT/EP 99/07055

| Patent document cited in search repo | n · | Publication date | Patent family member(s) | | | Publication date |
|---|-----|---------------------|-------------------------|----------|-------|---------------------|
| WO 9952938 | A | 21-10-1999 | DE | 19825585 | A | 21-10-1999 |
| , | ••• | | DE | 19828097 | Ā | 30-12-1999 |
| | | | DE | 19831637 | A | 27-01-2000 |
| | | | AŪ | 4120899 | Ā | 01-11-1999 |
| | | • | AŬ | 4481699 | A · | 01-11-1999 |
| | | | WO | 9952515 | Ā | 21-10-1999 |
| | | | WO | 9966875 | Ā | 29-12-1999 |
| | | | WO | 0004031 | | 27-01-2000 |
| | ٠ | | WO | 0003699 | A | 27-01-2000 |
| DE 19752700 | A | 02-06-1999 | DE | 29800547 | U | 08-04-1999 |
| | | 12 13 2000 | JP | 11169186 | Ă | 29-06-1999 |

Form PCT/ISA/210 (patent family annex) (July 1992)

hns. .tionelee Aldenzeichen PCT/EP 99/07055

| | | 101,2.00, | |
|---|--|---|--|
| A KLASSIF IPK 7 | FIZIERUNG DES ANMELDUNGSGEGENSTANDES C12N9/90 C12N9/10 C12N9/12 | C12Q1/48 | |
| Nach der Int | ernationsien Patentidassifikation (IPK) oder nach der nationalen Klass | stilkation und der IPK | |
| B. RECHER | CHIERTE GEBIETE | | |
| Recherchier IPK 7 | ter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbol C12N C12Q | e) | |
| Recherchier | te aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, son | welt diese unter die recherchierten Gebiete | fallen |
| Während de | r internationalen Recherche konsuttierte elektronische Dateribank (Na | ame der Datenbank und evti, verwendete & | Suchbegriffe). |
| C. ALS WE | SENTLICH ANGESEHENE UNTERLAGEN | | |
| Kategorie* | Bezeichnung der Veröffentlichung, soweit enforderlich unter Angabe | e der in Betracht kommenden Telle | Betr. Anspruch Nr. |
| E,L | WO 99 52938 A (HASSAN JOMAA) 21. Oktober 1999 (1999–10–21) | | 1,2,4, 8-12, 16-18 |
| | Siehe SeqID's; Priority of inv. sh W09952938 and PCTEP99/07055 and p DE19816196.4, DE19825585.3, DE19828 9831637.2, DE19831639.9 or DE19831 be invalid; A.4C(4) PC. | resent in 097.1,DE1 | |
| | | ·/ | |
| | tere Veröffentlichungen eind der Fontsetzung von Feld C zu nehmen | X Siehe Anhang Patentfamille | |
| "A" Veröfte aberr "E" äheres Anme "L" Veröfte schelt ander soll of suege "O" Veröfte eine E"P" Veröfte dern t | e Kategorien von angegebenen Veröffentlichungen : entlichung, die den allgemeinen Stand der Technik definiert, nicht als besonders bedeutsam anzuseinen ist Dokument, das jedoch erst am oder nach dem internationalen idedatum veröffentlicht worden ist entlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft er- nen zu lassen, oder durch die das Veröffentlichungsdatum einer ren im Recherchenbeheicht genannten Veröffentlichungsdatum einer ren im Recherchenbeheicht genannten Veröffentlichung belegt werden der die aus einem anderen besonderen Grund angegeben ist (wie ertlichung, die sich auf eine mündliche Offenbarung, Berutzung, eine Ausstellung oder andere Maßnahmen bezieht ertlichung, die vor dem internationalen Armeidedatum, über nach beanspruchten Prioritätsdatum veröffentlicht worden ist Abechtusses der internationalen Recherche | T' Spätere Veröffentlichung, die nach der oder dem Prioritätsdatum veröffentlich Armeldung nicht kolidiert, sondem nu Erfindung zugrundellegenden Prinzipa Theorie angegeben list "X" Veröffentlichung von besonderer Beder kann allein aufgrund dieser Veröffentlichung von besonderer Beder kann nicht als auf erfinderlicher Tätigkeit beruhrend betra "Y" Veröffentlichung von besonderer Beder kann nicht als auf erfinderlichung mit Veröffentlichung mit Veröffentlichung mit Veröffentlichung die Verbindung til einen Fachmann diese Veröffentlichung, die Nittglied derselber Absendedatum des Intermationalen Re | t worden let und mit der rzum Verständnie des der oder der ihr zugrundellegenden utung; die beanspruchte Erfindung chung nicht als neu oder auf achtet werden utung; die beanspruchte Erfindung reit beruhend betrachtet ietner oder mehreren anderen Verbindung gebracht wird und nahellegend list n Patentfamilie ist |
| 7 | 7. März 2000 | 23/03/2000 | |
| Name und | Postanschrift der Internationalen Recherchenbehörde Europäisches Patentarnt, P.B. 5816 Patentiaan 2 NL – 2280 HV Rijewijk Tel. (+51–70) 340–2040, Tx. 31 651 epo ni, | Bevolinächtigter Bediensteter | |
| 1 . | Fax: (+31-70) 340-3016 | l Hoekstra, S | |

Formblett PCT/ISA/210 (Blatt 2) (Juli 1992)

Inc. ationales Aldenzsichen
PCT/EP 99/07055

| etegorie* | ung) ALS WESENTLICH ANGESEHENE UNTERLAGEN Bezeichnung der Veröffentlichung, soweit erfordenlich unter Angabe der in Betracht kommenden Telle | Betr. Anspruch Nr. |
|-----------|--|--------------------|
| medoue, | nergen am in det Aetalietinetinist kohet etimpeleti mins Villane det til begrutt kolletjenet i ete | Seu, Ampiusi tu. |
| | PUTRA ET AL: "Incorporation of '2,3-'13!C2!- and '2,4-'13!C2!-D-1-Deoxyxylulose into ubiquinone of Escherichia coli via the Mevalonate-Independent pathway for Isoprenoid Biosynthesis" TETRAHEDRON LETTERS, NL, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, Bd. 39, Nr. 39, 1998, Seiten 23-26-26, XP002116676 ISSN: 0040-4039 Abbildung 1 | 15 |
| | KUZUYAMA ET AL: "Direct formation of 2-C-Methyl-D-Erythritol 4-phosphate from 1-Deoxy-D-Xylulose 5-phosphate Reductoisomerase, a new enzyme in the non-mevalonate pathway to isopentenyl diphosphate" TETRAHEDRON LETTERS, NL, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, Bd. 39, Nr. 39, 1998, Seiten 4509-4512-44512, XP002116675 ISSN: 0040-4039 Abbildung 1 | 15 |
| Ρ,Χ | SCHWENDER, J. ET AL.: "Cloning and heterologous expression of a cDNA encoding 1-deoxy-d-xylulose-5-phosphate reductoisomerase of Arabidopsis thaliana" FEBS LETTERS, Bd. 455, Juli 1999 (1999-07), Seiten 140-144, XP002132424 das ganze Dokument | 1,9-12 |
| Ρ,Α | DE 197 52 700 A (HOECHST SCHERING AGREVO GMBH) 2. Juni 1999 (1999-06-02) das ganze Dokument | 1-12 |
| A | LANGE ET AL: "A family of transketolases that directs isoprenoid biosynthesis via a mevalonate-independent pathway" FASEB JOURNAL, US, FED. OF AMERICAN SOC. FOR EXPERIMENTAL BIOLOGY, BETHESDA, MD, Bd. 95, März 1998 (1998-03), Seiten 2100-2104, XP002116672 ISSN: 0892-6638 das ganze Dokument | 1-18 |
| Ρ,Χ | EMINY DATABASE: "AC: EF111813" PLASMODIUM FALCIPARUM 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE, 11. Januar 1999 (1999-01-11), XP002132425 Siehe: Scores | 1,9-12 |
| | -/ | |

Int. Monelee Attenzeichen
PCT/EP 99/07055

| | | PCI/EP 99 | 07033 |
|-------------|--|-------------|--------------------|
| C.(Fortsetz | ing) ALS WESENTLICH ANGESEHENE UNTERLAGEN | | |
| Kategorie* | Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht komme | enden Telle | Betr. Anspruch Nr. |
| P,X | TREMBL DATABASE: "AC: 096693" PLASMODIUM FALCIPARUM 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE, 1. Mai 1999 (1999-05-01), XP002132426 Siehe: Scores | · | 1,9-12 |
| X | SPRENGER ET AL: "Identification of a thiamin-dependent synthase in Escherichia coli required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin, and pyridoxol" FASEB JOURNAL,US,FED. OF AMERICAN SOC. FOR EXPERIMENTAL BIOLOGY, BETHESDA, MD, Bd. 94, November 1997 (1997-11), Seiten 12857-12862, XPO02116674 ISSN: 0892-6638 Abbildung 2 | | 2,9-12 |
| A | Abbildung 1 | • | 15 |
| P,X | TREMBL DATABASE: "AC: QZ8HO" CHLAMYDIA PNEUMONIAE GCPE PROTEIN, 1. Mai 1999 (1999-05-01), XP002132427 Siehe: scores | | 3,9-12 |
| | | | |
| | | | |

Angaben zu Veröffentlichungen, die zur seiben Patentfamilie gehören

Int. Jonales Aktorizelohen
PCT/EP 99/07055

| Im Recherchenbericht, angeführtes Patentdokument | | Datum der Veröffentlichung | Mitglied(er) der Patentiamilie | | Datum der Veröffentlichung | |
|--|----------|--------------------------------------|-----------------------------------|----|-------------------------------|--------------------------------------|
| WO | 9952938 | A | 21-10-1999 | DE | 19825585 A | 21-10-1999 |
| | | | | DE | 198280 97 A | 30 - 12 - 19 99 |
| | • | | | DE | 198316 37 A | 27 - 01-200 0 |
| | | | • | AU | 4120899 A | 01-11-1999 |
| | | | Ť | AU | 4481699 A | 01-11-1999 |
| | | | | MO | 9952515 A | 21-10-1999 |
| | | | | WO | 9966875 A | 29-12-1999 |
| | | | | WO | 0004031 A | 27-01-2000 |
| | | | | MO | 0003699 A | 27-01-2000 |
| DE | 19752700 | A | 02-06-1999 | DE | 29800547 U | 08-04-1999 |
| | | | | JP | 11169186 A | 29-06-1999 |

Ren1 - 00-084 01/24335

N E R A C , Inc. One Technology Dr. Tolland, CT 06084 (860) 872-7000

DOCUMENT SERVICE

Order No. 0293339-001

This copy is being furnished for private research use only. It may not be further reproduced, resold, or used for publication. The customer assumes full responsibility for copyright questions that may arise concerning this reproduction or the use of the material. NERAC assumes no liability for contents of this material, nor use thereof.

TO:

Ms. Darlene Rentschler Renessen, LLC Suite 300 3000 Lakeside Drive Bannockburn, IL 60015

0293339-001

APR 1 5 2003 LEGAL DEPT.

SAVE THIS PACKING SLIP TO COMPARE WITH INVOICE AND WITH STATEMENT

CUST.ORDER DATE: 04/08/03 NERAC QUESTION:

ORDER RECEIVED: 04/08/03

DOCUMENT MAILED: 04/08/03 P.O. NO. 00-084

TITLE: Canadian Patent # CA2343919 Genes of the

1-desoxy-d-dxylulose biosynthetic pathway

LANGUAGE: English only

OPIC OFFICE DE LA PROPRIÉTÉ INTELLECTUELLE DU CANADA



(12)(19)(CA) Demande-Application

CIPO
CANADIAN INTELLECTUAL
PROPERTY OFFICE

(21)(A1) 2,343,919

(86) 1998/09/22 (87) 2000/03/30

- (72) JOMAA, HASSAN, DE
- (71) JOMAA, HASSAN, DE
- (51) Int.Cl.⁷ C12N 9/90, C12Q 1/48, C12N 9/12, C12N 9/10
- (30) 1998/09/22 (198 43 279.8) DE
- (30) 1999/05/21 (199 23 567.8) DE
- (54) VOIE DE SYNTHESE BIOLOGIQUE DES GENES DES 1-DESOXY-D-XYLULOSE
- (54) GENES OF THE 1-DESOXY-D-XYLULOSE BIOSYNTHETIC PATHWAY

(57) The invention relates to the 1-desoxy- D-xylulose-5-phosphate reductoisomerase gene, the 1-desoxy- D-xylulose-5-phosphate-synthase gene and the gcpE gene of the 1-desoxy- D-xylulose biosynthetic pathway and to their use for transforming vectors, host organisms and plants and for determining substances that inhibit this biosynthetic pathway.

PCT

WELTORGANISATION FÜR GEISTIGES EIGENTUM Internationales Büro

INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

| (51) Internationale Patentklassifikation 7: | | (11) Internationale Veröffentlichungsnummer: | WO 00/17233 |
|---|----|--|----------------------|
| C07K 14/00 | A2 | (43) Internationales | · |
| | ļ | | März 2000 (30.03.00) |

(21) Internationales Aktenzeichen:

PCT/EP99/07055

(22) Internationales Anmeldedatum:

22. September 1999

(22.09.99)

(30) Prioritätsdaten:

198 43 279.8 199 23 567.8 22. September 1998 (22.09.98) DE

21. Mai 1999 (21.05.99)

DE

(71)(72) Anmelder und Erfinder: JOMAA, Hassan [DE/DE]; Breslauer Strasse 24, D-35398 Gießen (DE).

(74) Anwälte: PANTEN, Kirsten usw.; Reichel und Reichel, Parkstrasse 13, D-60322 Frankfurt am Main (DE). (81) Bestimmungsstaaten: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), curasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), curopäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, TT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht

Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.

(54) Title: GENES OF THE 1-DESOXY-D-XYLULOSE BIOSYNTHETIC PATHWAY

(54) Bezeichnung: GENE DES 1-DESOXY-D-XYLULOSE-BIOSYNTHESEWEGS

(57) Abstract

The invention relates to the 1-desoxy- D-xylulose- 5-phosphate reductoisomerase gene, the 1-desoxy- D-xylulose- 5-phosphate-synthase gene and the gcpE gene of the 1-desoxy- D-xylulose biosynthetic pathway and to their use for transforming vectors, host organisms and plants and for determining substances that inhibit this biosynthetic pathway.

(57) Zusammenfassung

Die vorliegende Erfindung betrifft das 1-Desoxy- D-xylulose- 5-phosphatreduktoisomerase -Gen, das 1-Desoxy- D-xylulose- 5-phosphat- Synthase- Gen und das gcpE-Gen des 1-Desoxy- D-xylulose- Biosynthesewegs und ihre Verwendung zur Transformation von Vektoren, Wirtsorganismen und Pflanzen und zur Bestimmung von Stoffen, die diesen Biosyntheseweg inhibiteren.

-#-

Claims

- 1. DNA sequences which code for a polypeptide with the amino acid sequence shown in SEQ ID no. 2 or for an analogue or derivative of the polypeptide according to SEQ ID no. 2, in which one or more amino acids have been deleted, added or replaced by other amino acids, wherein the enzymatic action of the polypeptide is retained, and which sequences originate from parasites, wherein sequence variations occurring within the framework of natural strain variability are included.
- 2. DNA sequences which code for a polypeptide with the
 amino acid sequence shown in SEQ ID no. 4 or for an
 analogue or derivative of the polypeptide according
 to SEQ ID no. 4, in which one or more amino acids
 have been deleted, added or replaced by other amino
 acids, wherein the enzymatic action of the
 polypeptide is retained, and which sequences
 originate from parasites, wherein sequence
 variations occurring within the framework of natural
 strain variability are included.
- 25 3. DNA sequences which code for a polypeptide with the amino acid sequence shown in SEQ ID no. 6 or for an analogue or derivative of the polypeptide according to SEQ ID no. 6, in which one or more amino acids have been deleted, added or replaced by other amino acids wherein the catalytic function of the polypeptide is retained.

- 1 -

4. DNA sequence according to one of claims 1 to 3, characterised in that it also comprises functional regulation signals, in particular promoters, operators, enhancers, ribosomal binding sites.

5

- 5. DNA sequence with the following sub-sequences
 - i) promoter which is active in viruses, eukaryotes and prokaryotes and ensures the formation of an RNA in the intended target tissue or target cells,

10

- ii) DNA sequences according to one of claims 1 to 3,
- iii) 3' untranslated sequence which, in viruses, eukaryotes and prokaryotes, results in the addition of poly(A) residues onto the 3' end of the RNA.

15

20

6. Process for the production of transgenic viruses, eukaryotes and prokaryotes for modifying the isoprenoid content, characterised in that a DNA sequence according to claim 4 or 5 is transferred and incorporated into the genome of viruses, eukaryotic and prokaryotic cells with or without use of a vector.

25

7. Transgenic systems, in particular plants and plant cells which contain one or more DNA sequences according to claims 1 to 5 as "foreign" or "additional" DNA, which sequences are expressed.

30

8. Expression vector containing one or more DNA sequences according to claims 1 to 5.

PCT/EP99/07055

-il -

- 9. Protein which is involved in the 1-deoxy-D-xylulose 5-phosphate metabolic pathway and a) is coded by DNA sequences SEQ ID no. 1, 3 or 5 or b) is coded by DNA sequences which hybridise with DNA sequences SEQ ID no. 1, 3, 5 or fragments of these DNA sequences in the DNA region which codes for the mature protein.
- 10. Protein according to claim 9, obtainable from the culture supernatants of parasites or from the disrupted parasites and purification by chromatographic and electrophoretic methods.
- 11. Protein according to one of claims 9 and 10, characterised in that it a) is the product of viral,

 15 prokaryotic or eukaryotic expression of exogenous DNA, b) is coded by sequences SEQ ID no. 1, 3 or 5 or is coded by DNA sequences which hybridise with DNA sequences SEQ ID no. 1, 3, 5 or fragments of these DNA sequences in the DNA region which codes

 20 for the mature protein, or c) is coded by DNA sequences which would hybridise without degeneration of the genetic code with the sequences defined in b) and which code for a polypeptide with a corresponding amino acid sequence.

25

5

10

- 12. Protein according to one of the preceding claims, characterised in that it comprises the amino acid sequences SEQ ID no. 2, 4 or 6.
- 30 13. Process for determining the enzymatic activity of the gcpE protein, characterised in that phosphorylation of a sugar or of a phosphorus sugar or of a precursor of isoprenoid biosynthesis, in

WO 00/17233

5

-0-

particular the phosphorylation of 2-C-methyl-D-erythritol, 2-C-methyl-D-erytritol phosphate, in particular 2-C-methyl-D-erythritol 4-phosphate, 2-C-methyl-D-erythrose phosphate, in particular 2-C-methyl-D-erythrose 4-phosphate, and of phosphate and alcohol precursors, is detected.

14. Process according to claim 13, characterised in that 10 phosphorylation of the following phosphates or alcohols is detected: $CH_2(OH) - C(CH_3) = C(OH) - CH_2 - O - PO(OH)_2$ $CH_2(OH) - C(CH_3) = C(OH) - CH_2 - OH$ $CH_2(OH) - CH(CH_3) - CO - CH_2 - O - PO(OH)_2$ 15 CH₂ (OH) -CH (CH₃) -CO-CH₂OH $CH_2=C(CH_3)-CO-CH_2-O-PO(OH)_2$, $CH_2=C(CH_3)-CO-CH_2-OH$, $CH_2=C(CH_3)-CH(OH)-CH_2-O-PO(OH)_2$ $CH_2=C(CH_3)-CH(OH)-CH_2-OH$, 20 $CH_2(OH) - C(=CH_2) - C(OH) - CH_2 - O - PO(OH)_2$ $CH_2(OH) - C(=CH_2) - C(OH) - CH_2 - OH$ $CHO-CH(CH_3)-CH(OH)-CH_2-O-PO-(OH)_2$, $CHO-CH(CH_3)-CH(OH)-CH_2-OH$, $CH_2(OH) - C(OH)(CH_3) - CH = CH - O - PO(OH)_2$ 25 $CH_2(OH) - C(OH)(CH_3) - CH = CH - OH$ $CH(OH) = C(CH_3) - CH(OH) - CH_2 - O - PO(OH)_2$ $CH(OH) = C(CH_3) - CH(OH) - CH_2 - OH$ $(CH_3)_2HC-CO-CH_2-O-PO(OH)_2$ (CH₃)₂HC-CO-CH₂-O-H,30 $(CH_3)_2HC-CH(OH)-CH_2-O-PO(OH)_2$,

 $(CH_3)_2HC-CH(OH)-CH_2-O-H$.

20

30

- 1 -

- 15. Process for the combined determination of the enzymatic activity of DOXP synthase and of DOXP reductase, characterised in that the conversion of glyceraldehyde 3-phosphate into 2-C-methylerythritol 4-phosphate is detected.
- 16. Process for screening a compound for the treatment of infectious processes in humans and animals, wherein the process comprises:
- a) provision of a host cell which contains a recombinant expression vector, wherein the vector comprises at least a portion of the oligonucleotide sequence according to SEQ ID no. 1, SEQ ID no. 3 or SEQ ID no. 5 or variants or analogues thereof, and moreover of a compound suspected to have antimycotic, antibiotic, antiparasitic or antiviral action in humans and animals,
 - b) bringing the host cell into contact with the compound and
 - c) determining the antimicrobial, antimycotic, antibiotic, antiparasitic or antiviral action of the compound.
- 25 17. Process for screening for compounds for treating plants, wherein the process comprises:
 - a) provision of a host cell which contains a recombinant expression vector, wherein the vector comprises at least a portion of the oligonucleotide sequence according to SEQ ID no. 1, SEQ ID no. 3 or SEQ ID no. 5 or variants or analogues thereof, and moreover of a compound suspected to have antimicrobial,

- 9 -

- antiviral, antiparasitic, bactericidal, fungicidal or herbicidal action in plants,
- b) bringing the host cell into contact with the compound and
- c) determining the antimicrobial, antiviral, antiparasitic, bactericidal, fungicidal or herbicidal action of the compound.
- 18. Use of DNA according to one of claims 1 to 5 or of proteins according to one of claims 9 to 12 or of transgenic systems according to claim 7 for the prevention or treatment of diseases in humans and animals.

30

Genes of the 1-deoxy-D-xylulose biosynthesis pathway

The present invention relates to DNA sequences which, when incorporated into the genome of viruses, eukaryotes and prokaryotes, modify isoprenoid biosynthesis and to a genetic engineering process for the production of these transgenic viruses, eukaryotes and prokaryotes. The invention also relates to a process for the identification of substances having herbicidal, antimicrobial, antiparasitic, antiviral, fungicidal, bactericidal action in plants and antimicrobial, antiparasitic, antimycotic, antibacterial and antiviral action in humans and animals.

The biosynthesis pathway for the formation of isoprenoids via the classical acetate/mevalonate pathway and an alternative mevalonate-independent biosynthesis pathway, the deoxy-D-xylulose phosphate pathway is already known (Rohmer, M., Knani, M., Simonin, P., Sutter, B. and Sahm, H. (1993): Biochem. J. 295: 517-524).

It is, however, not known how and by which pathways it is possible to bring about a change in the isoprenoid concentration in viruses, eukaryotes and prokaryotes by means of the deoxy-D-xylulose phosphate pathway. Figure 1 shows this biosynthesis pathway.

DNA sequences are consequently provided which code for 1-deoxy-D-xylulase 5-phosphate synthase (DOXP synthase), 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DOXP reductoisomerase) or the gcpE protein. All three genes and enzymes are involved in isoprenoid biosynthesis.

-2-

(Translator's comment: The portion at the beginning of the next paragraph enclosed in square brackets corresponds to the beginning of the sentence which finishes on page 2, line 1 of the original).

[The gcpE protein has a kinase function and catalyses the phosphorylation of a sugar or a phosphorus sugar or a precursor of isoprenoid biosynthesis, in particular the phosphorylation of 2-C-methyl-D-erythritol, 2-C-methyl-D-erytritol phosphate, in particular 2-C-methyl-D-erythritol 4-phosphate, 2-C-methyl-D-erythrose, 2-C-

methyl-D-erythrose] phosphate, in particular 2-C-methyl-D-erythrose 4-phosphate. In the precursor of isoprenoid synthesis, the gcpE protein in particular catalyses the phosphorylation of the following substances:

 $CH_2(OH) - C(CH_3) = C(OH) - CH_2 - O - PO(OH)_2$,

15 CH₂ (OH) -C (CH₃) =C (OH) -CH₂-OH, CH₂ (OH) -CH (CH₃) -CO-CH₂-O-PO (OH)₂, CH₂ (OH) -CH (CH₃) -CO-CH₂OH CH₂=C (CH₃) -CO-CH₂-O-PO (OH)₂, CH₂=C (CH₃) -CO-CH₂-OH,

20 CH₂=C (CH₃) -CH (OH) -CH₂-O-PO (OH)₂,

CH₂=C (CH₃) -CH (OH) -CH₂-OH,

CH₂ (OH) -C (=CH₂) -C (OH) -CH₂-O-PO (OH)₂,

CH₂ (OH) -C (=CH₂) -C (OH) -CH₂-OH

CHO-CH (CH₃) -CH (OH) -CH₂-O-PO- (OH)₂,

25 CHO-CH (CH₃) -CH (OH) -CH₂-OH,

CH₂ (OH) -C (OH) (CH₃) -CH=CH-O-PO (OH)₂,

CH₂ (OH) -C (OH) (CH₃) -CH=CH-OH

CH (OH) =C (CH₃) -CH (OH) -CH₂-O-PO (OH)₂,

CH (OH) =C (CH₃) -CH (OH) -CH₂-OH,

30 (CH₃)₂HC-CO-CH₂-O-PO (OH)₂, (CH₃)₂HC-CO-CH₂-O-H, (CH₃)₂HC-CH (OH) -CH₂-O-PO (OH)₂, (CH₃)₂HC-CH (OH) -CH₂-O-H.

10

15

20

25

30

-25

DOXP synthase catalyses the condensation of pyruvate and glyceraldehyde 3-phosphate to yield 1-deoxy-D-xylulose 5-phosphate and DOXP reductoisomerase catalyses the conversion of 1-deoxy-D-xylulose 5-phosphate into 2-C-methyl-D-erythritol 4-phosphate (c.f. Fig. 1).

The invention relates to the following DNA sequences:

DNA sequences which code for a polypeptide with the amino acid sequence shown in SEQ ID no. 2 or for an analogue or derivative of the polypeptide according to SEQ ID no. 2, in which one or more amino acids have been deleted, added or replaced by other amino acids, wherein the enzymatic action of the polypeptide is retained, and which sequences originate from parasites, wherein sequence variations occurring within the framework of natural strain variability are included,

DNA sequences which code for a polypeptide with the amino acid sequence shown in SEQ ID no. 4 or for an analogue or derivative of the polypeptide according to SEQ ID no. 4, in which one or more amino acids have been deleted, added or replaced by other amino acids, wherein the enzymatic action of the polypeptide is retained, and which sequences originate from parasites, wherein sequence variations occurring within the framework of natural strain variability are included,

and DNA sequences which code for a polypeptide with the amino acid sequence shown in SEQ ID no. 6 or for an analogue or derivative of the polypeptide according to SEQ ID no. 6, in which one or more amino acids have been

- 3 -

deleted, added or replaced by other amino acids, wherein the catalytic function of the polypeptide is retained.

The genes and the gene products thereof (polypeptides)

are shown with their primary structure and are assigned as follows:

SEQ ID no. 1: 1-deoxy-D-xylulose 5-phosphate reductoisomerase gene

SEQ ID no. 2: 1-deoxy-D-xylulose 5-phosphate reductoisomerase

SEQ ID no. 3: 1-deoxy-D-xylulose 5-phosphate synthase gene

SEQ ID no. 4: 1-deoxy-D-xylulose 5-phosphate synthase

- 4 -

PCT/EP99/07055

SEQ ID no. 5: gcpE gene

5

10

SEQ ID no. 6: gcpE proteins.

The DNA sequences all originate from *Plasmodium* falciparum.

Apart from the DNA sequences stated in the sequence listing, suitable sequences are also those which, as a result of the degeneration of the genetic code, have another DNA sequence, but code for the same peptide or for an analogue or derivative of the polypeptide, in which one or more amino acids have been deleted, added or replaced by other amino acids.

- The sequences according to the invention are suitable for the expression of genes in viruses, eukaryotes and prokaryotes which are responsible for isoprenoid biosynthesis in the 1-deoxy-D-xylulose pathway.
- According to the invention, eukaryotes or eukaryotic cells include animal cells, plant cells, algae, yeasts, fungi, while prokaryotes or prokaryotic cells include bacteria, archaebacteria and eubacteria.
- When a DNA sequence is incorporated into a genome on which the above-stated DNA sequence is located, expression of the above-described genes in viruses, eukaryotes and prokaryotes is enabled. The viruses, eukaryotes and prokaryotes transformed according to the invention are cultivated in a manner known per se and the isoprenoid formed during such cultivation is isolated and optionally purified. Not all isoprenoids need to be

15

20

25

isolated as in some case the isoprenoids are releas d directly into the ambient air.

The invention furthermore relates to a process for the production of transgenic viruses, eukaryotes and prokaryotes in order to modify the isoprenoid content, which process comprises the following steps.

- a) Production of a DNA sequence with the following sub sequences
 - promoter which is active in viruses, eukaryotes and prokaryotes and ensures the formation of an RNA in the intended target tissue or target cells,
 - ii) DNA sequence which codes for a polypeptide with the amino acid sequence shown in SEQ ID no. 2, 4 or 6 or for an analogue or derivative of the polypeptide according to SEQ ID no. 2, 4 or 6,
 - iii) 5' and 3' untranslated sequence which enables or enhances expression of the stated genes in viruses, eukaryotes and prokaryotes,
 - b) transfer and incorporation of the DNA sequence into the genome of viruses, prokaryotic or eukaryotic cells with or without the use of a vector (for example plasmid, viral DNA).

The intact, whole plants may be regenerated from plant cells transformed in this manner.

The protein-coding sequences with the nucleotide sequences SEQ ID no. 1, SEQ ID no. 3 and SEQ ID no. 5 may be provided with a promoter which ensures transcription in certain organs or cells, which promoter is coupled in

- 6 -

sense orientation (3' end of the promoter to the 5' end of the coding sequence) to the sequence which codes the protein to be formed. A termination signal which determines termination of mRNA synthesis is attached to the 3' end of the coding sequence. In order to direct the 5 protein which is to be expressed to certain subcellular compartments, such as chloroplasts, amyloplasts, mitochondria, vacuoles, cytosol or intercellular spaces, a further sequence which codes for a so-called signal 10 sequence or a transit peptide may be inserted between the promoter and the coding sequence. In some cases, it is necessary to insert sequences which code for a signal at the COOH terminus of the protein. The sequence must be in the same reading frame as the coding sequence of the 15 protein. A large number of cloning vectors is available in order to prepare for the introduction of the DNA sequences according to the invention into higher plants, which vectors contain a replication signal for E. coli and a marker which permits selection of the transformed cells. Depending upon the method by which desired genes 20 are introduced into the plant, further DNA sequences may be required. If, for example, the Ti or Ri plasmid is used to transform the plant cells, at least one right border, but frequently the right border and left border 25 of the Ti and Ri plasmid T-DNA must be inserted as a flanking region into the genes to be introduced. The use of T-DNA for transforming plant cells has been intensively investigated and comprehensively described in EP 120516; Hoekama in "The Binary Plant Vector System", 30 Offset-drukkerij Kanters B.V. Alblasserdam (1985), chapter V; Fraley et al., Crit.Rev.Plant Sci. 4, 1-46 and An et al. (1985) EMBO J. 4, 277-287. Once the introduced DNA has been incorporated into the genome, it is

30

generally stable and is also retained in the descendants of the originally transformed cells. It normally contains a selection marker, which imparts to the transformed plant cells resistance to a biocide or an antibiotic, such as kanamycin, G 418, bleomycin, hygromycin or phosphinotricin and others. The particular marker used is thus intended to allow selection of transformed cells from cells lacking the inserted DNA.

10 Many techniques are available for introducing DNA into a plant. These techniques include transformation with the assistance of agrobacteria, for example Agrobacterium tumefaciens, protoplast fusion, microinjection of DNA, electroporation, as well as ballistic methods and virus infection. Whole plants may then be regenerated from the 15 transformed plant material in a suitable medium which may contain antibiotics or biocides for selection purposes. No particular requirements are placed upon the plasmids for injection and electroporation. However, if whole plants are to be regenerated from such transformed cells, 20 a selectable marker gene must be present. The transformed cells grow in the plants in the conventional manner (McCormick et al. (1986), Plant Cell Reports 5, 81-84). The plants may be cultivated normally and be crossed with plants which have the same transformed genome or other 25 genomes. The resultant individuals have the corresponding phenotypic properties.

The present invention also provides expression vectors which contain one or more of the DNA sequences according to the invention. Such expression vectors are obtained by providing the DNA sequences according to the invention with suitable functional regulation signals. Such

regulation signals are DNA sequences which are responsible for expression, for example promoters, operators, enhancers, ribosomal binding sites, and are recognised by the host organism.

5

Further regulation signals, which for example control replication or recombination of the recombinant DNA in the host organism, may optionally also be a constituent part of the expression vector.

10

The host organisms transformed with the DNA sequences or expression vectors according to the invention are also provided by the present invention.

15 Suitable host cells and organisms for expressing the enzymes according to the invention are those which comprise no intrinsic enzymes with the function of DOXP synthase, DOXP reductoisomerase or the gcpE protein. This is the case for archaebacteria, animals, fungi, slime 20 moulds and some eubacteria. The absence of such intrinsic enzyme activity substantially facilitates detection and purification of the recombinant enzymes. As a consequence, it is also for the first time possible straightforwardly to measure, in crude extracts from the 25 host cells, the activity and in particular the inhibition of the activity of the recombinant enzymes according to the invention by various chemicals and pharmaceuticals.

The enzymes according to the invention are advantageously
then expressed in eukaryotic cells if post-translational
modification and native folding of the polypeptide chain
is to be achieved. Moreover, depending upon the
expression system, it is ensured when expressing genomic

DNA sequences that introns are eliminated by splicing the DNA and the enzymes are produced in the polypeptide sequences characteristic to the parasites. Using recombinant DNA techniques, sequences coding for introns may be eliminated from or inserted for experimental purposes into the DNA sequences to be expressed.

The protein may be isolated from the host cell or the culture supernatant of the host cell using methods known to the person skilled in the art. *In vitro* reactivation of the enzymes may also be required.

In order to facilitate purification, the enzymes according to the invention or sub-sequences of the enzymes may be expressed as fusion proteins with various peptide chains. Oligo-histidine sequences and sequences derived from glutathione S-transferase, thioredoxin or calmodulin-binding peptides are particularly suitable for this purpose.

20

25

5

10

15

The enzymes according to the invention or sub-sequences of the enzymes may furthermore be expressed as fusion proteins with such peptide chains known to the person skilled in the art that the recombinant enzymes are transported into the extracellular medium or into certain compartments of the host cells. Both purification and investigation of the biological activity of the enzymes may consequently be facilitated.

When expressing the enzymes according to the invention, it may prove convenient to modify individual codons.

Purposeful replacement of bases in the coding region may here also be advisable if the codons used in the

parasites differ from the codon use in the heterologous expression system, in order to ensure optimal synthesis of the protein.

The enzymes according to the invention may furthermore be obtained under standardised conditions by in vitro translation by methods known to the person skilled in the art. Systems suitable for this purpose are rabbit reticulocyte and wheat germ extracts and bacterial lysates. In vitro transcribed mRNA may also be translated into Xenopus oocytes.

Oligo- and polypeptides, the sequences of which are derived from the peptide sequence of the enzymes according to the invention, may be obtained by chemical synthesis. Given appropriate selection of the sequences, such peptides have properties which are characteristic of the enzymes according to the invention. Such peptides may be produced in large quantities and are particularly suitable for investigating the kinetics of enzyme activity, regulation of enzyme activity, the three-dimensional structure of the enzymes, inhibition of enzyme activity by various chemicals and pharmaceuticals and the binding geometry and binding affinity of various ligands.

DNA with the nucleotides from sequences SEQ ID no. 1, 3 and 5 are preferably used for the recombinant production of the enzymes according to the invention.

30

25

15

20

The invention accordingly moreover relates to a process for screening for compounds which inhibit the deoxy-D-xylulose phosphate metabolic pathway. According to this

5

10

30

process, a host organism, which contains a recombinant expression vector, wherein the vector comprises at least a portion of the oligonucleotide sequence according to SEQ ID no. 1, SEQ ID no. 3 or SEQ ID no. 5 or variants or homologues thereof, is provided, as is a compound which is suspected to have antimicrobial, antiparasitic, antibacterial, antiviral and antimycotic action in humans and animals or an antimicrobial, antiviral, bactericidal, herbicidal or fungicidal activity in plants. The host organism is then brought into contact with the compound and the activity of the compound determined.

The present invention also provides methods for determining the enzymatic activity of the gcpE protein. 15 Said activity may be determined using known methods. Determination is performed by detecting the phosphorylation of a sugar or of a phosphorus sugar or of a precursor of isoprenoid biosynthesis, in particular the phosphorylation of 2-C-methyl-D-erythritol, 2-C-methyl-D-20 erytritol phosphate, in particular 2-C-methyl-Derythritol 4-phosphate, 2-C-methyl-D-erythrose, 2-Cmethyl-D-erythrose phosphate, in particular 2-C-methyl-Derythrose 4-phosphate. The present invention also provides the use of this measurement method for 25 identifying substances which inhibit the activity of the particular enzymes.

The enzymatic activity of DOXP synthase and DOXP reductoisomerase may be detected in a single step by determining the conversion of glyceraldehyde 3-phosphate into 2-C-methylerythritol 4-phosphate.

5

- 12 -

Determination of the activities of DOXP synthase and DOXP reductoisomerase proceeds analogously. Fluorimetric methods described by Querol et al. are also suitable for determining DOXP synthase activity (Querol et al., abstracts, 4th European Symposium on Plant Isoprenoids, Barcelona, 21-23 April 1999).

-1-

SEQUENCE LISTING

```
<110> Jomaa, Hassan
<120> Genes of the 1-deoxy-D-xylulose biosynthesis pathway
     <130> 15696
     <140> PCT/EP99
     <141> 1999-09-22
     <150> DE19923567.8
     <151> 1999-05-22
     <150> DE19843279.8
     <151> 1998-09-22
     <160> 6
     <170> PatentIn Ver. 2.1
     <210> 1 '
     <211> 1467
     <212> DNA
     <213> Plasmodium falciparum
     <220>
     <221> CDS
     <222> (1)..(1467)
     <220>
     <221> gene
     <222> (1)..(1467)
     <220>
     <221>. mRNA
     <222> (1)..(1467) '
     <400> 1
     Met Lys Lys Tyr Ile Tyr Ile Tyr Phe Phe Ile Thr Ile Thr Ile
     aat gat tta gta ata aat aca tca aaa tgt gtt tcc att gaa aga
                                                                  96
     Asn Asp Leu Val Ile Asn Asn Thr Ser Lys Cys Val Ser Ile Glu Arg
     aga aaa aat aac gca tat ata aat tat ggt ata gga tat aat gga cca
     Arg Lys Asn Asn Ala Tyr Ile Asn Tyr Gly Ile Gly Tyr Asn Gly Pro
     gat aat aaa ata aca aag agt aga aga tgt aaa aga ata aag tta tgc
     Asp Asn Lys Ile Thr Lys Ser Arg Arg Cys Lys Arg Ile Lys Leu Cys
```

| aaa Lys 65 | aag Lys | gat Asp | tta Leu | ata Ile | gat Asp 70 | att Ile | ggt Gly | gca Ala | ata Ile | aag Lys 75 | aaa Lys | cca Pro | att Ile | aat Asn | gta Val 80 | 240 |
|------------------|------------|------------|------------|------------|------------------|-------------------|------------|------------|------------|------------------|------------|------------|------------|------------|------------------|-----|
| _ | | | | - | | ggt Gly | _ | | | _ | | | | | | 288 |
| | | | | | | att Ile | | | | | | | | | | 336 |
| | | | _ | - | | aat Asn | - | | | - | | _ | _ | - | | 384 |
| | | | | | | ata Ile 135 | | | | | | | | | | 432 |
| | - | | _ | | | ata Ile | | - | | | | | | _ | - | 480 |
| | - | - | | _ | | gaa Glu | | _ | - | _ | | - | | - | | 528 |
| | - | | | | - | tct Ser | | | | | | | | _ | | 576 |
| | | | | | | ata Ile | | | | | | | | | | 624 |
| _ | | - | | | | tta Leu 215 | - | | | | | | | | | 672 |
| Ala 225 | Lys | Ile | Ile | Pro | Val 230 | gat Asp | Ser | Glu | His | Ser 235 | Ala | Ile | Phe | Gln | Cys 240 | 720 |
| | _ | | | _ | _ | tta Leu | | | | _ | | | _ | | | 768 |
| | | | | | | aat Asn | | | | | | | | | | 816 |
| | | | | | | | | | | | | | | | | |

4.30 **

| | - | | _ | | | | | | _ | | - | | ata Ile | | | 912 |
|---|---|---|---|-------------------|---|---|---|-----|---|---|-----|---|-------------------|---|---|------|
| - | | _ | | _ | _ | | | | | | _ | | gaa Glu | | | 960 |
| | | | - | - | _ | | | _ | | _ | - | | gta Val | | | 1008 |
| | | | | | | | | | | | | | tca Ser 350 | | | 1056 |
| _ | | - | | | | - | _ | | 4 | | | | tat Tyr | | | 1104 |
| | | | | | | | | | | | | | gat Asp | | | 1152 |
| _ | - | | | | | | | | | | | _ | cat His | | _ | 1200 |
| | | | | | | | | | | | | | ttt Phe | | | 1248 |
| | - | | | | | | - | | - | | | | ttt Phe 430 | - | | 1296 |
| | | | | | | _ | | | | | Ile | _ | caa Gln | _ | | 1344 |
| _ | | | | | | _ | _ | | - | | - | - | gat Asp | | - | 1392 |
| _ | | | | | | | | | - | | _ | | gct Ala | | - | 1440 |
| | | | | cat His 485 | | | | tag | | | | | | | • | 1467 |

<210> 2

<211> 488 <212> PRT <213> Plasmodium falciparum

<400> 2 Met Lys Lys Tyr Ile Tyr Ile Tyr Phe Phe Phe Ile Thr Ile Thr Ile Asn Asp Leu Val Ile Asn Asn Thr Ser Lys Cys Val Ser Ile Glu Arg Arg Lys Asn Asn Ala Tyr Ile Asn Tyr Gly Ile Gly Tyr Asn Gly Pro Asp Asn Lys Ile Thr Lys Ser Arg Arg Cys Lys Arg Ile Lys Leu Cys Lys Lys Asp Leu Ile Asp Ile Gly Ala Ile Lys Lys Pro Ile Asn Val Ala Ile Phe Gly Ser Thr Gly Ser Ile Gly Thr Asn Ala Leu Asn Ile Ile Arg Glu Cys Asn Lys Ile Glu Asn Val Phe Asn Val Lys Ala Leu Tyr Val Asn Lys Ser Val Asn Glu Leu Tyr Glu Gln Ala Arg Glu Phe 120 Leu Pro Glu Tyr Leu Cys Ile His Asp Lys Ser Val Tyr Glu Glu Leu 135 Lys Glu Leu Val Lys Asn Ile Lys Asp Tyr Lys Pro Ile Ile Leu Cys Gly Asp Glu Gly Met Lys Glu Ile Cys Ser Ser Asn Ser Ile Asp Lys Ile Val Ile Gly Ile Asp Ser Phe Gln Gly Leu Tyr Ser Thr Met Tyr Ala Ile Met Asn Asn Lys Ile Val Ala Leu Ala Asn Lys Glu Ser Ile 200 Val Ser Ala Gly Phe Phe Leu Lys Lys Leu Leu Asn Ile His Lys Asn Ala Lys Ile Ile Pro Val Asp Ser Glu His Ser Ala Ile Phe Gln Cys 230 235 Leu Asp Asn Asn Lys Val Leu Lys Thr Lys Cys Leu Gln Asp Asn Phe Ser Lys Ile Asn Asn Ile Asn Lys Ile Phe Leu Cys Ser Ser Gly Gly 265 Pro Phe Gln Asn Leu Thr Met Asp Glu Leu Lys Asn Val Thr Ser Glu Asn Ala Leu Lys His Pro Lys Trp Lys Met Gly Lys Lys Ile Thr Ile

Asp S r Ala Thr Met Met Asn Lys Gly Leu Glu Val Ile Glu Thr His 305 310 315 320

Phe Leu Phe Asp Val Asp Tyr Asn Asp Ile Glu Val Ile Val His Lys 325 330 335

Glu Cys Ile Ile His Ser Cys Val Glu Phe Ile Asp Lys Ser Val Ile 340 345 350

Ser Gln Met Tyr Tyr Pro Asp Met Gln Ile Pro Ile Leu Tyr Ser Leu 355 360 365

Thr Trp Pro Asp Arg Ile Lys Thr Asn Leu Lys Pro Leu Asp Leu Ala 370 375 380

Gln Val Ser Thr Leu Thr Phe His Lys Pro Ser Leu Glu His Phe Pro 385 390 395 400

Cys Ile Lys Leu Ala Tyr Gln Ala Gly Ile Lys Gly Asn Phe Tyr Pro 405 410 415

Thr Val Leu Asn Ala Ser Asn Glu Ile Ala Asn Asn Leu Phe Leu Asn 420 425 430

Asn Lys Ile Lys Tyr Phe Asp Ile Ser Ser Ile Ile Ser Gln Val Leu 435 440 445

Glu Ser Phe Asn Ser Gln Lys Val Ser Glu Asn Ser Glu Asp Leu Met 450 455 460

Lys Gln Ile Leu Gln Ile His Ser Trp Ala Lys Asp Lys Ala Thr Asp 465 470 475 480

Ile Tyr Asn Lys His Asn Ser Ser . 485

<210> 3

<211> 3872

<212> DNA

<213> Plasmodium falciparum

<220>

<221> CDS

<222> (126)..(3740)

<220>

<221> gene

<222> (1)..(3870)

<220>

<221> mRNA

<222> (1)..(3870)

<400> 3

ggtaatatac gtataatata tatataatat attcttacgt atgtatcatt tatgaatcat 60

| aata | aatat | ttc t | aaaţ | tta | c tt | ccgt | tttt | gct | cgat | tctt | ctca | attti | tcg 1 | tttc | agcttt | 120 |
|-----------|-------|-------|------|-----|------|------|------|-----|------|-------|------|-------|-------|-------------------|------------------------|-----|
| tate | | | | | | | | | | ys As | | | | | tt gtt al Val 15 | 170 |
| | | | | | | | | | | | | | _ | aat Asn 30 | | 218 |
| | | | | | | - | | | | | _ | | - | aat Asn | | 266 |
| _ | | | | | _ | | _ | _ | | | | | | gca Ala | | 314 |
| - | | - | | | | - | - | | _ | | - | | | ggc Gly | | 362 |
| | | | | | | | | | | | | | | aat Asn | | 410 |
| | _ | | | - | | _ | - | _ | | | | | | aat Asn 110 | _ | 458 |
| att Il | _ | | | | | | | | | | | | _ | agt Ser | - | 506 |
| | | | - | | _ | | | _ | | | | - | - | tta Leu | | 554 |
| | | | | | | | | | | | | | | gct Ala | | 602 |
| | | | | | | | | | | | | | | ata Ile | | 650 |
| | | | | | | | | | | | | | | gac Asp 190 | | 698 |
| | | | _ | | _ | | | - | | | _ | | - | gcc Ala | | 746 |
| | | | | - | | | | | | _ | | | - | aaa Lys | | 794 |

-7-

| cgt Arg 225 | | | | | | | | | | | 842 |
|-----------------------|---|---|--|---|---|---|---|---|---|---|------|
| aat Asn | | | | | | | | | | | 890 |
| aat Asn | | | | | | | | | | | 938 |
| gat Asp | | | | | | | | | | | 986 |
| aga Arg | | | | | | | | | | | 1034 |
| aat Asn 305 | | | | | | | | | | | 1082 |
| aat Asn | | | | - | - | | | | | _ | 1130 |
| ata Ile | | | | | | | | | | | 1178 |
| tat Tyr | | | | | | | | | | | 1226 |
| ttt Phe | _ | - | | | _ | - | _ | - | - | | 1274 |
| gaa Glu 385 | _ | | | | _ | | | | - | | 1322 |
| tat Tyr | | | | | | | | | | | 1370 |
| ttc Phe | | | | | | | | - | | _ | 1418 |
| aaa Lys | | | | | | | | | | | 1466 |

-8-

| | | | | - | | ata Ile | | | | | | | | - | | 1514 |
|--|---|---|--|--|--|---|---|--|---|---|---|--|---|---|---|----------------------|
| agc Ser | | | | | | tta Leu 470 | | | | | | | | | | 1562 |
| | | | | | | gat Asp | | | | | | | | | | 1610 |
| | _ | | | - | | cta Leu | | | | | | - | | | | 1658 |
| | | | | | | aat Asn | | | | | | | | | | 1706 |
| | | | | | | act Thr | | | | | | | | | | 1754 |
| - | - | | | | | aag Lys 550 | | | _ | | | | | | _ | 1802 |
| ata | caa | 345 | -~+ | | | | | | | _ | | | | | | |
| | - | | _ | _ | | gca Ala | | - | _ | | aat Asn | _ | | | | 1850 |
| Ile 560 caa | Glu aaa | Ile gga | Ser | Asp | Asn 565 aat | _ | Asn aat | Val aat | Thr | Asn 570 aac | Asn aat | Glu aat | Arg | Ile aat | Phe 575 aat | 1850 |
| Ile 560 caa Gln aat | Glu aaa Lys aat | Ile gga Gly tat | Ser ata Ile | Asp cac His 580 aat | Asn 565 aat Asn cct | Ala gat | Asn aat Asn gat Asp | Val aat Asn gtg | Thr att Ile 585 gta | Asn 570 aac Asn gga | Asn aat Asn aga | Glu aat Asn gaa | Arg att Ile aat | aat Asn 590 | Phe 575 aat Asn | |
| Ile 560 caa Gln aat Asn | Glu aaa Lys aat Asn | gga Gly tat Tyr | Ser ata Ile atc Ile 595 gta | Asp cac His 580 aat Asn | Asn 565 aat Asn cct Pro | Ala gat Asp | aat Asn gat Asp | Val aat Asn gtg Val 600 | Thr att Ile 585 gta Val | Asn 570 aac Asn gga Gly | Asn aat Asn aga Arg | Glu aat Asn gaa Glu aaa | att Ile aat Asn 605 gta | aat Asn 590 acg Thr | Phe 575 aat Asn aat Asn | 1898 |
| Ile 560 caa Gln aat Asn gta Val | Glu aaa Lys aat Asn cca Pro | gga Gly tat Tyr aat Asn 610 | ser ata Ile atc Ile 595 gta Val | Asp cac His 580 aat Asn cga Arg | Asn 565 aat Asn cct Pro aat Asn | Ala gat Asp tca Ser | Asn aat Asn gat Asp aac Asn 615 | Val aat Asn gtg Val 600 cat His | Thr att Ile 585 gta Val aac Asn | Asn 570 aac Asn gga Gly gtg Val | Asn aat Asn aga Arg gat Asp | Glu aat Asn gaa Glu aaa Lys 620 gca Ala | att Ile aat Asn 605 gta Val | aat Asn 590 acg Thr cac | Phe 575 aat Asn aat Asn att Ile | 1898 |
| Ile 560 caa Gln aat Asn gta Val gct Ala tta | Glu aaa Lys aat Asn cca Pro att Ile 625 aat | gga Gly tat Tyr aat Asn 610 ata Ile | ser ata Ile atc Ile 595 gta Val gga Gly | Asp cac His 580 aat Asn cga Arg gat Asp | Asn 565 aat Asn cct Pro aat Asn ggt Gly | Ala gat Asp tca Ser gat Asp | aat Asn gat Asp aac Asn 615 tta Leu | Val aat Asn gtg Val 600 cat His aca Thr | Thr att Ile 585 gta Val aac Asn ggt Gly | Asn 570 aac Asn gga Gly gtg Val gga Gly att | Asn aat Asn aga Arg gat Asp atg Met 635 | Glu aat Asn gaa Glu aaa Lys 620 gca Ala | att Ile aat Asn 605 gta Val tta Leu | aat Asn 590 acg Thr cac His gaa Glu | Phe 575 aat Asn aat Asn att Ile gcg Ala aat | 1898 1946 1994 |

WO 00/17233 PCT/EP99/07055

-9-

| | - | | | | | | | _ | | tta Leu | | | | • | | 2186 |
|------------|------------|-------------------|------------|------------|------------|------------|-------------------|------------|------------|-------------------|------------|-------------------|------------|------------|------------|------|
| | | - | _ | | - | | - | | | tta Leu | _ | | | _ | | 2234 |
| | | | | | | | | | | gat Asp | | | | | | 2282 |
| | | | | | | | | | | gta Val 730 | | | | | | 2330 |
| | | | | | | | | | | cat His | | | | | | 2378 |
| _ | | - | | | | | _ | _ | | ata Ile | - | | _ | | | 2426 |
| | | | | | | | | | | acc Thr | | | | | | 2474 |
| | | | - | | | _ | | _ | _ | gag Glu | | | | | | 2522 |
| | | | | | | | | | | aat Asn 810 | | | | | | 2570 |
| | | | | | | | | | | tca Ser | | | | | | 2618 |
| | | | | | | | | | | tta Leu | | | | | | 2666 |
| ata Ile | ata Ile | ttc Phe 850 | cta Leu | tct Ser | ccc Pro | gct Ala | atg Met 855 | tta Leu | gga Gly | gga Gly | tca Ser | gga Gly 860 | ttg Leu | gtt Val | aaa Lys | 2714 |
| | | | | | | | | | | gat Asp | | | | | | 2762 |
| | | | | | | | | | | gca Ala 890 | | | | | | 2810 |

4)

| aaa Lys | ata Ile | caa Gln | tta Leu | tgt Cys 900 | ata Ile | tat Tyr | tcg Ser | acc Thr | ttt Phe 905 | tta Leu | caa Gln | aga Arg | gca Ala | tat Tyr 910 | gat Asp | 2858 |
|------------|-------------------|--------------------|-------------------|--------------------|------------|-------------------|--------------------|--------------------|-------------------|-------------|-------------------|--------------------|--------------------|-------------------|------------|------|
| | | | | gat Asp | | | | | | | | | | | ata Ile | 2906 |
| att Ile | Ğĺy | aga Arg 930 | agt Ser | gga Gly | tta Leu | gta Val | gga Gly 935 | gag Glu | gat Asp | G] À aàà | gca Ala | aca Thr 940 | cat His | caa Gln | ggt Gly | 2954 |
| ata Ile | tat Tyr 945 | gat Asp | tta Leu | tct Ser | tat Tyr | ctt Leu 950 | ggg ggg | aca Thr | ctt Leu | aac Asn | aat Asn 955 | gca Ala | tat Tyr | ata Ile | ata Ile | 3002 |
| | | | | caa Gln | | | | | | | | | | | | 3050 |
| tta Leu | gat Asp | aag Lys | gac Asp | cat His 980 | tct Ser | gtg Val | tat Tyr | ata Ile | cgt Arg 985 | ata Ile | ccc Pro | aga Arg | atg Met | aac Asn 990 | ata Ile | 3098 |
| tta L u | agt Ser | gat Asp | aag Lys 995 | tac Tyr | atg Met | aaa Lys | Gly | tat Tyr 1000 | ttg Leu | aac Asn | att Ile | His | atg Met 1005 | aaa Lys | aat Asn | 3146 |
| gag Glu | Ser | aaa Lys 1010 | Asn | atc Ile | gat Asp | Val | aac Asn 1015 | gtg Val | gat Asp | ata Ile | Asn | gat Asp 1020 | gat Asp | gta Val | gat Asp | 3194 |
| Lys | | Ser | | gaa Glu | Tyr | | | | | Asn | | | | | | 3242 |
| | Gly | | | aga Arg | | | | | Asp | | | | | Asn | | 3290 |
| | | | Tyr | tca Ser 1060 | | | | Asp | | | | | Lys | | | 3338 |
| | | Ile | | aac Asn | | | Ser | | | | | Val | | | | 3366 |
| | Lys | | Ile | gaa Glu | | Glu | | | | | His | | | | | 3434 |
| Ser | | Val | | atg Met | Ile | | Leu | | | Leu | | Lys | | | | 3482 |

- 11 -

| | cat His) | | | Lys | | | | | Gln | | | | | Tyr | | 3530 |
|------------------|---------------------------------|-----------|-------------------|------------|------------|------------|-----------|--------------------|------------|------------|------------|-----------|--------------------|-----------|------------|------|
| | aat Asn | | 11e | | | | | Thr | | | | | Tyr | | | 3578 |
| gaa Glu | aat Asn | Asn | tat Tyr 155 | att Ile | aca Thr | aaa Lys | His | aac Asn 1160 | tta Leu | tat Tyr | gtt Val | His | aat Asn 1165 | Ile | tat Tyr | 3626 |
| | tct Ser | | | | | Ğlu | | - | | | Lys | - | | | - | 3674 |
| Val | gtc Val 1185 | | | | Lys | | | | | Asn | | | | | | 3722 |
| | aaa Lys) | | | Pro | | tgat | tgtaa | aga t | aaat | catai | ta ti | ttct | aaaa | t | | 3770 |
| tati | tttt | tt t | tata | actti | ta at | tgtgt | tacaa | a taa | aaata | atat | atc | taaai | tat a | attt | tatttg | 3830 |
| tac | gcttt | ctt t | tttt | tttt | tt t | taat | ttgti | t att | ttt | gtat | at | | | | | 3872 |
| <213 <213 | 0> 4 L> 12 2> PE B> PI | RT | odiur | n fal | lcipa | arum | | | | | | | | | | |
| <400 Met 1 |)> 4 Ile | Phe | Asn | Tyr 5 | Val | Phe | Phe | Lys | Asn 10 | Phe | Val | Pro | Val | Val 15 | Leu | |
| Tyr | Ile | Ŀeu | Leu 20 | Ile | Ile | Tyr | Ile | Asn 25 | Leu | Asn | Gly | Met | Asn 30 | Asn | Lys | |
| Asn | Gln | Ile 35 | Lys | Thr | Glu | Lys | Ile 40 | Tyr | Ile | Lys | Lys | Leu 45 | Asn | Arg | Leu | |
| Ser | Arg 50 | Lys | Asn | Ser | Leu | Cys. 55 | Ser | Ser | Lys | Asn | Lys 60 | Ile | Ala | Cys | Leu | |
| Phe 65 | Asp | Ile | Gly | Asn | Asp 70 | Asp | Asn | Arg | Asn | Thr 75 | Thr | Tyr | Gly | Tyr | Asn 80 | |
| Val | Asn | Val | Lys | Asn 85 | Asp | Asp | Ile | Asn | Ser 90 | Leu | Leu | Lys | Asn | Asn 95 | Tyr | |
| Ser | Asn | Lys | Leu | Tyr | Met | Asp | Lys | Arg | Lys | Asn | Ile | Asn | Asn | Val | Ile | |
| | | | 100 | - | | _ | | 105 | - | | | | 110 | | | |

Gln Lys Glu Asn Glu Gln Lys Arg Asn Lys Gln Arg Cys Leu Thr Gln Cys His Thr Tyr Asn Met Ser His Glu Gln Asp Lys Leu Ala Asn Asp 150 155 Asn Asn Arg Asn Asn Lys Lys Asn Phe Asn Leu Leu Phe Ile Asn Tyr . 170 Phe Asn Leu Lys Arg Met Lys Asn Ser Leu Leu Asn Lys Asp Asn Phe 185 Phe Tyr Cys Lys Glu Lys Lys Leu Ser Phe Leu His Lys Ala Tyr Lys Lys Lys Asn Cys Thr Phe Gln Asn Tyr Ser Leu Lys Arg Lys Ser Asn Arg Asp Ser His Lys Leu Phe Ser Gly Glu Phe Asp Asp Tyr Thr Asn Asn Asn Ala Leu Tyr Glu Ser Glu Lys Lys Glu Tyr Ile Thr Leu Asn 250 Asn Asn Asn Lys Asn Asn Asn Asn Asn Asn Asn Asp Asn Lys Asn Asn 265 Asp Asn Asn Asp Tyr Asn Asn Asn Ser Cys Asn Asn Leu Gly Glu Arg Ser Asn His Tyr Asp Asn Tyr Gly Gly Asp Asn Asn Asn Pro Cys 295 Asn Asn Asn Asn Asp Lys Tyr Asp Ile Gly Lys Tyr Phe Lys Gln Ile Asn Thr Phe Ile Asn Ile Asp Glu Tyr Lys Thr Ile Tyr Gly Asp Glu 325 330 Ile Tyr Lys Glu Ile Tyr Glu Leu Tyr Val Glu Arg Asn Ile Pro Glu Tyr Tyr Glu Arg Lys Tyr Phe Ser Glu Asp Ile Lys Lys Ser Val Leu Phe Asp Ile Asp Lys Tyr Asn Asp Val Glu Phe Glu Lys Ala Ile Lys Glu Glu Phe Ile Asn Asn Gly Val Tyr Ile Asn Asn Ile Asp Asn Thr Tyr Tyr Lys Lys Glu Asn Ile Leu Ile Met Lys Lys Ile Leu His Tyr 410 Phe Pro Leu Leu Lys Leu Ile Asn Asn Pro Ser Asp Leu Lys Lys Leu 430

Lys Lys Gln Tyr Leu Pro Leu Leu Ala His Glu Leu Lys Ile Phe Leu Phe Phe Ile Val Asn Ile Thr Gly Gly His Phe Ser Ser Val Leu Ser Ser Leu Glu Ile Gln Leu Leu Leu Tyr Ile Phe Asn Gln Pro Tyr Asp Asn Val Ile Tyr Asp Ile Gly His Gln Ala Tyr Val His Lys Ile 485 Leu Thr Gly Arg Lys Leu Leu Phe Leu Ser Leu Arg Asn Lys Lys Gly Ile Ser Gly Phe Leu Asn Ile Phe Glu Ser Ile Tyr Asp Lys Phe Gly Ala Gly His Ser Ser Thr Ser Leu Ser Ala Ile Gln Gly Tyr Tyr Glu 535 Ala Glu Trp Gln Val Lys Asn Lys Glu Lys Tyr Gly Asn Gly Asp Ile Glu Ile Ser Asp Asn Ala Asn Val Thr Asn Asn Glu Arg Ile Phe Gln 565 570 Lys Gly Ile His Asn Asp Asn Asn Ile Asn Asn Ile Asn Asn Asn 585 Asn Tyr Ile Asn Pro Ser Asp Val Val Gly Arg Glu Asn Thr Asn Val Pro Asn Val Arg Asn Asp Asn His Asn Val Asp Lys Val His Ile Ala 615 Ile Ile Gly Asp Gly Gly Leu Thr Gly Gly Met Ala Leu Glu Ala Leu 625 630 635 Asn Tyr Ile Ser Phe Leu Asn Ser Lys Ile Leu Ile Ile Tyr Asn Asp Asn Gly Gln Val Ser Leu Pro Thr Asn Ala Val Ser Ile Ser Gly Asn Arg Pro Ile Gly Ser Ile Ser Asp His Leu His Tyr Phe Val Ser Asn 680 Ile Glu Ala Asn Ala Gly Asp Asn Lys Leu Ser Lys Asn Ala Lys Glu Asn Asn Ile Phe Glu Asn Leu Asn Tyr Asp Tyr Ile Gly Val Val Asn 715 Gly Asn Asn Thr Glu Glu Leu Phe Lys Val Leu Asn Asn Ile Lys Glu

Asn Lys Leu Lys Arg Ala Thr Val Leu His Val Arg Thr Lys Lys S r Asn Asp Phe Ile Asn Ser Lys Ser Pro Ile Ser Ile Leu His Ser Ile 760 Lys Lys Asn Glu Ile Phe Pro Phe Asp Thr Thr Ile Leu Asn Gly Asn Ile His Lys Glu Asn Lys Ile Glu Glu Glu Lys Asn Val Ser Ser Ser Thr Lys Tyr Asp Val Asn Asn Lys Asn Asn Lys Asn Asn Asp Asn Ser Glu Ile Ile Lys Tyr Glu Asp Met Phe Ser Lys Glu Thr Phe Thr Asp 825 820 Ile Tyr Thr Asn Glu Met Leu Lys Tyr Leu Lys Lys Asp Arg Asn Ile Ile Phe Leu Ser Pro Ala Met Leu Gly Gly Ser Gly Leu Val Lys Ile Ser Glu Arg Tyr Pro Asn Asn Val Tyr Asp Val Gly Ile Ala Glu Gln His Ser Val Thr Phe Ala Ala Ala Met Ala Met Asn Lys Lys Leu Lys 885 Ile Gln Leu Cys Ile Tyr Ser Thr Phe Leu Gln Arg Ala Tyr Asp Gln Ile Ile His Asp Leu Asn Leu Gln Asn Ile Pro Leu Lys Val Ile Ile Gly Arg Ser Gly Leu Val Gly Glu Asp Gly Ala Thr His Gln Gly Ile 930 935 Tyr Asp Leu Ser Tyr Leu Gly Thr Leu Asn Asn Ala Tyr Ile Ile Ser Pro Ser Asn Gln Val Asp Leu Lys Arg Ala Leu Arg Phe Ala Tyr Leu Asp Lys Asp His Ser Val Tyr Ile Arg Ile Pro Arg Met Asn Ile Leu Ser Asp Lys Tyr Met Lys Gly Tyr Leu Asn Ile His Met Lys Asn Glu 1000 1005 Ser Lys Asn Ile Asp Val Asn Val Asp Ile Asn Asp Asp Val Asp Lys 1015 Tyr Ser Glu Glu Tyr Met Asp Asp Asp Asn Phe Ile Lys Ser Phe Ile 025 1035 1030

WO 00/17233 PCT/EP99/07055

- 15 -

Gly Lys Ser Arg Ile Ile Lys Met Asp Asn Glu Asn Asn Asn Thr Asn 1045 1050 1055

Glu His Tyr Ser Ser Arg Gly Asp Thr Gln Thr Lys Lys Lys Val 1060 1065 1070

Cys Ile Phe Asn Met Gly Ser Met Leu Phe Asn Val Ile Asn Ala Ile 1075 1080 1085

Lys Glu Ile Glu Lys Glu Gln Tyr Ile Ser His Asn Tyr Ser Phe Ser 1090 1095 1100

Ile Val Asp Met Ile Phe Leu Asn Pro Leu Asp Lys Asn Met Ile Asp 105 1110 1115 1120

His Val Ile Lys Gln Asn Lys His Gln Tyr Leu Ile Thr Tyr Glu Asp 1125 1130 1135

Asn Thr Ile Gly Gly Phe Ser Thr His Phe Asn Asn Tyr Leu Ile Glu 1140 1145 1150

Asn Asn Tyr Ile Thr Lys His Asn Leu Tyr Val His Asn Ile Tyr Leu 1155 1160 1165

Ser Asn Glu Pro Ile Glu His Ala Ser Phe Lys Asp Gln Glu Val 1170 1175 1180

Val Lys Met Asp Lys Cys Ser Leu Val Asn Arg Ile Lys Asn Tyr Leu 185 1190 1195 1200

Lys Asn Asn Pro Thr 1205

<210> 5

<211> 3147

<212> DNA

<213> Plasmodium falciparum

<220>

<221> CDS

<222> (199)..(2670)

<400> 5

tgcctgaata accacaaa atg agt tat ata aaa aga ctg att ctt ttt atg 231

Met Ser Tyr Ile Lys Arg Leu Ile Leu Phe Met

1 5 10

tta ctg ttt tat tct cat gta aaa att aaa aaa tta ttt att aaa att 279 L u Leu Ph Tyr Ser His Val Lys Ile Lys Lys Leu Ph Ile Lys Ile 15 20 25 - 16 -

| | aat Asn | | | | | | | | | 327 |
|---|-------------------|---|---|--|--|---|-----|---|---|-----|
| - | Phe 45 | | | | | | | - | _ | 375 |
| | act Thr | | | | | | | | | 423 |
| | tat Tyr | | | | | | | | | 471 |
| | aat Asn | | | | | | | | | 519 |
| | tat Tyr | | | | | | | | | 567 |
| | gtt Val 125 | | | | | | | | | 615 |
| | act Thr | | | | | | Glu | | | 663 |
| | att Ile | | | | | | | | | 711 |
| | caa Gln | | | | | | | | | 759 |
| | tta Leu | - | - | | | _ | - | • | | 807 |
| | cct Pro 205 | | | | | | | | | 855 |
| | aat Asn | | | | | | | | | 903 |
| | tat Tyr | | | | | | | | | 951 |

- 17 -

| - | | | _ | | | | - | | - | | aga Arg | | | _ | - | 999 |
|------------|-------------------|-------------------|------------|-------------------|------------|-------------------|-------------------|------------|-------------------|------------|-------------------|-------------------|------------|-------------------|------------|------|
| | | | | | | | | | | | tct Ser | | | | | 1047 |
| | | | | | | | | | | | tcg Ser 295 | | | | | 1095 |
| | | | | | | | | | | | ctt Leu | | | | | 1143 |
| | | | | | | | | | | | tat Tyr | | | | | 1191 |
| | | | | | | | | | | | ata Ile | | | | | 1239 |
| | | | | | | | | | | | aaa Lys | | | | | 1287 |
| | | | | | | | | | | | acc Thr 375 | | | | | 1335 |
| | | | | | | | | | | | tgt Cys | | | | | 1383 |
| gaa Glu | aat Asn | tta Leu | Lys | aaa Lys 400 | aga Arg | ata Ile | ttt Phe | tat Tyr | aat Asn 405 | gaa Glu | aat Asn | ttt Phe | aaa Lys | gaa Glu 410 | gat Asp | 1431 |
| | | | | | | | | | | | aat Asn | | | | | 1479 |
| gaa Glu | gaa Glu | aat Asn 430 | tat Tyr | cga Arg | aat Asn | ttt Phe | aat Asn 435 | aat Asn | ata Ile | aaa Lys | aaa Lys | aga Arg 440 | aat Asn | gta Val | gaa Glu | 1527 |
| aaa Lys | aat Asn 445 | aat Asn | aat Asn | gta Val | tta Leu | cat His 450 | gaa Glu | gag Glu | tgc Cys | act Thr | ata Ile 455 | ggt Gly | aat Asn | gta Val | gta Val | 1575 |
| | | | | | | | | | | | ttt Phe | | | | | 1623 |

- 18 -

| | | | | aat Asn | | | | | | | | 1671 |
|---|---|--|---|-------------------|---|---|---|---|---|---|---|------|
| | | | | aat Asn | | | | | | | | 1719 |
| | | | | tta Leu | | | | | | | | 1767 |
| | - | | | ata Ile | _ | | _ | _ | - | | | 1815 |
| _ | | | | aat Asn 545 | | | | | | | _ | 1863 |
| | | | _ | aac Asn | - | | | | | _ | - | 1911 |
| | | | | tta Leu | | | | - | - | | | 1959 |
| | | | | gaa Glu | | _ | - | | | | | 2007 |
| | | | | tat Tyr | | | | | | | | 2055 |
| | | | | tta Leu 625 | | | | | | | | 2103 |
| | | | | aat Asn | | | | | | | | 2151 |
| | | | | ttg Leu | | | | | | | | 2199 |
| | | | | gtt Val | | | | | | | | 2247 |
| | | | | gaa Glu | | | | | | | | 2295 |

WO 00/17233 PCT/EP99/07055

- 19 -

| | | | | | | | | | | | | | | | tta Leu | | 2343 |
|----------------------------|------------|------------|------------|------------|-------------------|------------|------------|------------|------------|-------------------|------------|------------|------------|------------|-------------------|------------|------|
| | | | | | | | | | | | | | | | ata Ile 730 | | 2391 |
| | | | | | | | | | | | | | | | act Thr | | 2439 |
| | | | | | | | | | | | | | | | gca Ala | | 2487 |
| Marient (f.) Grand (f.) | Met | | | | | | | | | | | | | | cat His | | 2535 |
| | | | | | | | | | | | | | | | ggt Gly | | 2583 |
| | gag Glu | tta Leu | gta Val | gaa Glu | aga Arg 800 | aat Asn | ata Ile | cct Pro | gag Glu | gaa Glu 805 | gaa Glu | gct Ala | tgt Cys | gat Asp | aaa Lys 810 | ttg Leu | 2631 |
| | | | | | | | | aac Asn | | | | | | taaa | ittga | at | 2680 |
| | atgg | acaa | igt a | attta | ttta | t tt | attt | atct | tat | atat | aat | atat | tata | aa t | tttt | cgatg | 2740 |
| | tatt | tttc | ct t | ttaa | aatt | t ta | tttt | tttt | tta | tttt | ttt | tttt | gaag | jta a | tatt | tataa | 2800 |
| | tgca | taca | ata a | atatt | aaaa | t gt | gtat | tata | taa | taat | atc | attt | tatt | gt t | attt | taaaa | 2860 |
| | gact | aata | icc -a | agaa | caat | t tt | ttaa | taat | cat | tctt | ata | actt | gtta | aa t | atat | atata | 2920 |
| | tata | tata | ta t | attt | attt | a tt | tata | ttta | tat | ttat | tta | tttt | tggt | at a | tgaa | aagta | 2980 |
| | aaaa | tata | at a | attt | aaaa | ıg ta | ttta | caaa | ata | aata | ata | ttat | atat | ct ç | tttt | tatat | 3040 |
| | atat | gtta | at ç | gaaa | agga | g aa | aata | aata | aat | aaaa | caa | acaa | aata | ac a | tata | tatat | 3100 |
| | atat | atat | at a | ctga | atga | g aa | agaa | aaaa | aaa | agaa | aag | gata | cga | | | | 3147 |

<210> 6

<211> 824

<212> PRT

<213> Plasmodium falciparum

<400> 6

Met Ser Tyr Ile Lys Arg Leu Ile L u Ph Met Leu Leu Phe Tyr Ser 1 5 10 15

His Val Lys Ile Lys Leu Phe Ile Lys Ile Ser Asn Val Asn Ile Phe Phe Ala Glu Ala Lys Lys Asn Gly Lys Lys Glu Phe Phe Leu Phe Leu Leu Asn Ile Lys Lys Asn Ser Gln Gln Lys Lys Thr Tyr His Ile Thr Lys Arg Asn Thr Ile Asn Lys Ser Asp Phe Leu Tyr Ser Leu Leu Asn Glu Glu Gly Asn Ser Ser Lys Lys Glu Tyr Lys Asn Leu Lys Asp Glu Glu Lys Tyr Asn Ile Ile Gln Asn Ile Lys Lys Tyr Cys Glu Cys Thr Lys Lys Tyr Lys Arg Leu Pro Thr Arg Glu Val Val Ile Gly Asn 120 Val Lys Ile Gly Gly Asn Asn Lys Ile Ala Ile Gln Thr Met Ala Ser 135 Cys Asp Thr Arg Asn Val Glu Glu Cys Val Tyr Gln Ile Arg Lys Cys Lys Asp Leu Gly Ala Asp Ile Val Arg Leu Thr Val Gln Gly Val Gln Glu Ala Gln Ala Ser Tyr His Ile Lys Glu Lys Leu Leu Ser Glu Asn 180 185 Val Asn Ile Pro Leu Val Ala Asp Ile His Phe Asn Pro Lys Ile Ala Leu Met Ala Ala Asp Val Phe Glu Lys Ile Arg Val Asn Pro Gly Asn Tyr Val Asp Gly Arg Lys Lys Trp Ile Asp Lys Val Tyr Lys Thr Lys Glu Glu Phe Asp Glu Gly Lys Leu Phe Ile Lys Glu Lys Phe Val Pro L u Ile Glu Lys Cys Lys Arg Leu Asn Arg Ala Ile Arg Ile Gly Thr Asn His Gly Ser Leu Ser Ser Arg Val Leu Ser Tyr Tyr Gly Asp Thr 280 Pro Leu Gly Met Val Glu Ser Ala Phe Glu Phe Ser Asp Leu Cys Ile 295 Glu Asn Asn Phe Tyr Asn Leu Val Phe Ser Met Lys Ala Ser Asn Ala 310 315 Tyr Val Met Ile Gln Ser Tyr Arg Leu Leu Val Ser Lys Gln Tyr Glu

- 21 -

| | | | | 325 | | | | | 330 | | | | | 335 | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Arg | Asn | Met | Met 340 | Phe | Pro | Ile | His | Leu 345 | Gly | Val | Thr | Glu | Ala 350 | Gly | Phe |
| G1 y | Asp | Asn 355 | Gly | Arg | lle | Lys | Ser 360 | Tyr | Leu | Gly | Ile | Gly 365 | Ser | ·Leu | Leu |
| Tyr | Asp 370 | Gly | Ile | Gly | Asp | Thr 375 | Ile | Arg | Ile | Ser | Leu 380 | Thr | Glu | Asp | Pro |
| Trp 385 | Glu | Glu | Leu | Thr | Pro 390 | Cys | Lys | Lys | Leu | Val 395 | Glu | Asn | Leu | Lys | Lys 400 |
| Arg | Ile | Phe | Tyr | Asn 405 | Glu | Asn | Phe | Lys | Glu 410 | Asp | Asn | Glu | Leu | Lys 415 | Asn |
| Asn | Glu | Met | Asp 420 | Thr | Lys | Asn | Leu | Leu 425 | Asn | Phe | Glu | Glu | Asn 430 | Tyr | Arg |
| Asn | Phe | Asn 435 | Asn | Ile | Lys | Lys | Arg 440 | Asn | Val | Glu | Lys | Asn 445 | Asn | Asn | Val |
| Leu | His 450 | Glu | Glu | Cys | Thr | Ile 455 | Gly | Asn | Val | Val | Thr 460 | Ile | Lys | Glu | Leu |
| Glu 465 | Asp | Ser | Leu | Gln | Ile 470 | Phe | Lys | Asp | Leu | Asn 475 | Leu | Glu | Val | Asp | Ser 480 |
| Asn | Gly | Asn | Leu | Lys 485 | Lys | Gly | Ala | Lys | Thr 490 | Thr | Asp | Met | Val | Ile 495 | Ile |
| Asn | Asp | Phe | His 500 | Asn | Ile | Ťhr | Asn | Leu 505 | Gly | Lys | Lys | Thr | Val 510 | Asp | Lys |
| Leu | Met | Gln 515 | Val | Gly | Ile | Asn | Ile 520 | Val | Val | Ģln | Tyr | Glu 525 | Pro | His | Asn |
| Ile | Glu 530 | Phe | Ile | Glu | Lys | Met 535 | -Glu | Pro | Asn | Asn | Asp 540 | Asn | Asn | Asn | Asn |
| Asn 545 | Asn | Asn | Asn | Asn | Ile 550 | Leu | Phe | Tyr | Val | Asp 555 | Ile | Lys | Asn | Ile | Met 560 |
| Asn | Ser | Ser | Glu | Lys 565 | Asn | Ile | Lys | Leu | Ser 570 | Asn | Ser | Lys | Gly | Tyr 575 | Gly |
| Leu | Ile | Leu | Asn 580 | Gly | Lys | Glu | Asp | Ile 585 | Gln | Thr | Ile | Lys | Lys 590 | Ile | Lys |
| Glu | Leu | Asn 595 | Arg | Arg | Pro | Leu | Phe 600 | Ile | Leu | Leu | Lys | Ser 605 | Asp | Asn | Ile |
| Tyr | Glu 610 | His | Val | Leu | Ile | Thr 615 | Arg | Arg | Ile | Asn | Glu 620 | Leu | Leu | Gln | Ser |
| Leu 625 | Asn | Ile | Asn | Ile | Pro 630 | Tyr | Ile | His | Tyr | Val 635 | Asp | Ile | Asn | Ser | Asn 640 |

 Asn
 Tyr
 Asp
 Asp
 Leu 645
 Leu Val
 Asn
 Ser 650
 Leu Tyr
 Ala Gly 665
 Ser Cys 655

 Leu Met
 Asp 660
 Met 660
 Asp 660
 Asp 660
 Leu Met 665
 Leu 11e Val Asn Val Thr Asp 670
 Asp 670
 Asp Asp 670
 Asp 670
 Asp 670
 Asp 670
 Asp 681
 Asp 685
 Glu Thr Lys Tyr Asp 688
 Glu Lys Glu Lys Glu Glu Glu Glu Glu Asp 6880
 Asp 11e Lys Tyr 700
 Tyr Asp 11e His Arg Leu Lys Glu Glu Asp 710
 Asp 11e Lys Asp 11e Lys Thr 700
 Asp 710
 Asp 710</

Lys His Asn Lys Trp Lys Asp Pro 820